access

NCSA | Vol. 19, No. 3 | Fall 2006

access

Vol. 19 No. 3 Fall 2006

Executive Editor
J. William Bell
jbell@ncsa.uiuc.edu

Managing Editor
Barbara Jewett
barbaraj@ncsa.uiuc.edu

Art Director
Blake Harvey
bharvey@ncsa.uiuc.edu

Copyeditor
Kathleen Ricker

On the Web
News from NCSA is always available on the Web. Sign up for our RSS feed or email service for the latest developments from NCSA and its partners.
http://www.ncsa.uiuc.edu

National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
1205 West Clark Street
Urbana, IL 61801
217-244-0072

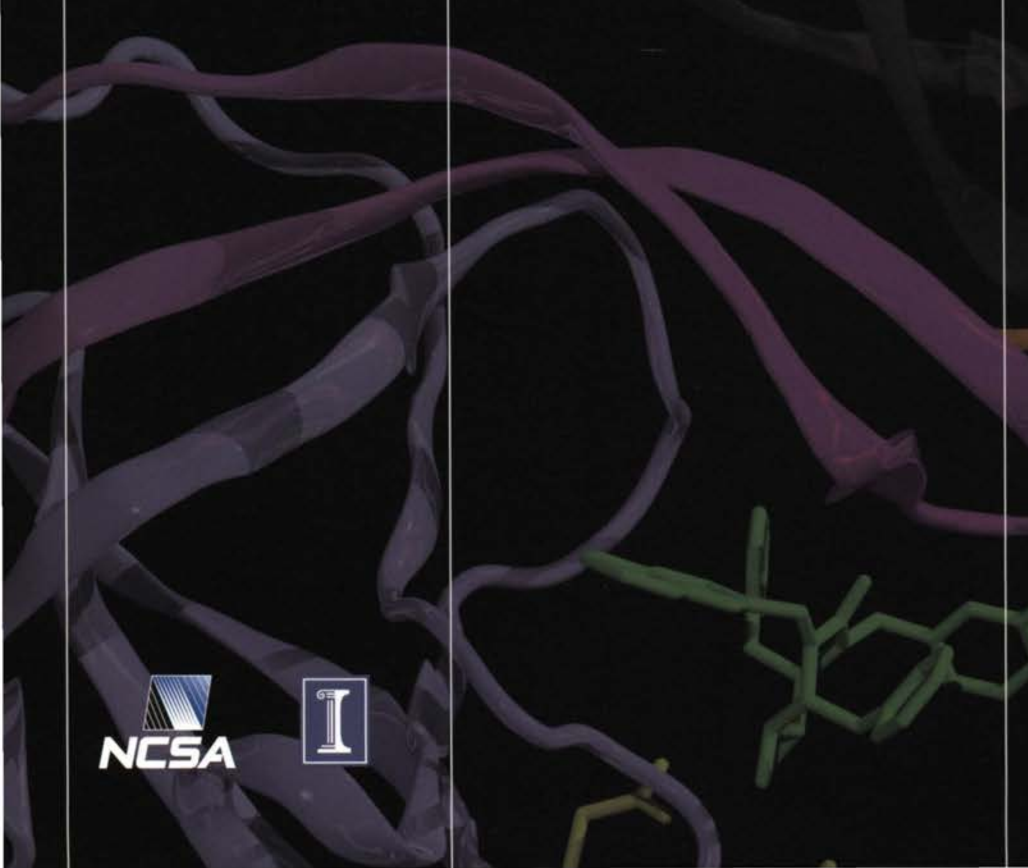
Who we are
The University of Illinois at Urbana-Champaign's National Center for Supercomputing Applications (NCSA), one of the five original centers in the National Science Foundation's Supercomputer Centers Program, opened its doors in January 1986. Over the past twenty years, NCSA has contributed significantly to the birth and growth of the worldwide cyberinfrastructure for science and engineering, operating some of the world's most powerful supercomputers and developing the software infrastructure needed to efficiently use them. Today the center is recognized as an international leader in deploying robust high-performance computing resources and in working with research communities to develop the new computing and software technologies needed to take full advantage of the rapidly expanding national and international cyberinfrastructure.

The center focuses on three themes. Cyberenvironments will give research communities the means to fully exploit the extraordinary resources available on the Internet (computing systems, data sources and stores, and tools). NCSA's cyber-resources ensure that computing, data, and networking resources are available to solve the most demanding science and engineering problems and that the solutions are obtained in a timely manner. Finally, innovative systems research explores the path to petascale computing—testing and evaluating the performance of emerging computing systems for scientific and engineering applications.

NCSA is a key partner in the National Science Foundation's TeraGrid project, a $100 million effort to offer researchers remote access to some of the fastest unclassified supercomputers as well as an unparalleled array of visualization tools, application software, sensors and instruments, and mass storage devices. NCSA also leads the effort to develop a secure national cyberinfrastructure through the National Center for Advanced Secure Systems Research, a project funded by the Office of Naval Research.

The center leaves its mark through the development of networking, visualization, storage, data management, data mining, and collaboration software as well. The prime example of this influence is NCSA Mosaic, which was the first graphical Web browser widely available to the general public. NCSA visualizations, meanwhile, have been a part of productions by the likes of PBS's NOVA and the Discovery Channel. Through its Private Sector Program, top researchers explore the newest hardware and software, virtual prototyping, visualization, networking, and data mining to help U.S. industries maintain a competitive edge in the global economy.

Major support for NCSA is provided by the National Science Foundation. Additional funding comes from the state of Illinois, industrial partners, and other federal agencies. For more information, see **http://www.ncsa.uiuc.edu/**.



On the cover
A team from Stony Brook University is working to develop a greater understanding of the mechanistic events associated with binding of HIV-1 protease substrates and inhibitors, which is critical for the design of more potent inhibitors of the enzyme. HIV protease is a dimer, formed when two identical chains come together (right and left).One of the simulations the team conducted using NCSA's Cobalt supercomputing system was to see what would happen when an inhibitor is added to the open structure.

In this simulation, an inhibitor (green) is manually placed near the entry to the active site of an open structure simulation without an inhibitor. The other flap closes and helps to pack the bulky side groups of the inhibitor into the pocket. One observation the team made during the simulation was that in structures with an inhibitor, the left flap (purple) is in front when the flaps close, while in structures without an inhibitor the right flap (orange) is in front.

NCSA

contents

Bob Wilhelmson
NCSA Chief Science Officer
page 02

Addressing Complexity through Synergy

An Expert Opinion

When I entered graduate school at the University of Illinois in the late 1960s, I was excited about working with a group on the development of Illiac IV, one of the first parallel computers. As I reflect on my experiences since, I recall many surprises, challenges, and successes associated with changes in high-performance architectures, the introduction of cyberenvironments, and advancements in numerical storm models.

The use of increasingly complex models has reproduced—sometimes with uncanny accuracy—such phenomena as supercells, with their mesocyclones and tornadoes, as well as severe weather associated with convective lines including bow echoes, severe squall lines, downbursts, and microbursts. Today, computational power has increased to the point that these convective models are being used in forecasting severe weather events. The return on the time and resources that have been spent is impressive, and the promise of more to come is exciting.

The years have also shown me that synergy is a crucial part both of generating simulations of increasing complexity and of exploiting that complexity. Today this not only consists of scientists working together but often consists of computer scientists, application and algorithm researchers, software developers, discipline scientists, service providers, educators, and students. An example is the LEAD (Linked Environments for Atmospheric Research) project funded by NSF involving nine institutions, including NCSA. LEAD is developing a comprehensive national cyberinfrastructure for mesoscale meteorology. It tackles the fundamental information technology research needed to create an integrated, scalable environment for identifying, accessing, preparing, assimilating, predicting, managing, analyzing, mining, and visualizing a broad array of meteorological data and model output. In other words, it's building a cyberenvironment—an integrated, end-to-end system that allows people to coordinate, automate, and apply high-end resources and capabilities.

Accomplishing this task requires a broad set of people, directed at lowering the barrier to using complex weather technologies. By democratizing advanced weather technologies, which traditionally were limited to government and university groups, the environment will empower a growing set of users. This facilitates rapid understanding, experiment, design, and execution of problem-solving simulations by both scientists and educators. The development of cyberenvironments, both within LEAD and at NCSA, encourage this increased human synergy and the increased human understanding that comes with it.

LEAD and NCSA are also focused on human-computer synergy. Efforts in dynamic adaptation are focused on responding automatically and in a coordinated manner to internal and external influences, often requiring on-demand access to computing resources. Dynamic adaptation to the weather includes the ability for numerical models and hazardous weather detection systems to respond to their data, driving new observations and launching new model simulations. Traditionally this been done statically and with much human intervention. Observations are taken in pre-selected patterns and model simulations are made at fixed time intervals and over fixed regions. But hazardous weather doesn't behave in a regular pattern or at regular times. And storm prediction has plenty of other avenues and issues that humans can focus on beyond setting up these simulations.

There have been many advances in individual technologies and model capabilities over the past 40 years that have increased the complexity of our ever-growing knowledge base. The increasing role of synergy in the development of models and cyberenvironments—along with their use by researchers, developers, forecasters, planners, and policymakers—enables us and many others in different disciplines to obtain new understanding and to continue to address important societal problems.

Bob Wilhelmson
Chief Science Officer
NCSA

Good CARMA

With support from the National Science Foundation, CARMA (the Combined Array for Research in Millimeter-wave Astronomy) joins telescopes from two earlier millimeter arrays—the Berkeley-Illinois-Maryland Association array and Caltech's Owens Valley Radio Observatory array—to form a more powerful astronomical tool. A grand opening for the high-altitude array was held in the spring, and NCSA's Trish Barker recently spoke with CARMA Director Anneila I. Sargent, the Benjamin M. Rosen Professor of Astronomy at Caltech, about the new opportunities CARMA will provide to observe and understand galaxies, molecular clouds forming clusters of stars, newly born stars emerging from their clouds, comets, and the cosmic radiation leftover from the Big Bang.

Q: Can you describe what CARMA is and why it is significant?

A: It's a combination of existing telescopes/antennas from Caltech's Owens Valley Radio Observatory array and the Berkeley-Illinois-Maryland Association (BIMA) array. It comprises at the moment six 10-meter telescopes from the Owens Valley array and nine 6-meter telescopes from the BIMA array.

I'd really like to emphasize that this is a novel and innovative array; it's not just the sum of some old parts put together. The changes we've introduced in terms of new technology and software make this a state-of-the-art instrument.

Having this new array with the mix of telescope sizes on a higher altitude site means that we get a combination of high sensitivity, high resolution, and the capability to image astronomical sources over wide fields of view at these millimeter wavelengths with better fidelity than has been achieved anywhere else.

Q: Many people probably think of a telescope as something that captures visible light, but these telescopes are looking at a different wavelength.

A: These are radio telescopes. They detect radiation at millimeter wavelengths. We're very sensitive to thermal radiation from forming stars, forming planets from the gas of the interstellar medium, and from other galaxies both near and far.

Often the most interesting regions of the universe—those where stars and planets are forming or the earliest epochs when galaxies are in their earliest stages of evolution—are obscured from optical view by intervening dust. These dust particles are minute, often comparable in size to the wavelength of visible light. As a result, they can prevent optical radiation from astronomical sources from reaching you directly. At millimeter wavelengths you can detect the emission from interstellar gas molecules and dust directly, and also infer what is happening behind the dust.

Q: How do the multiple telescopes in the array work together?

A: The size of a telescope is related to the degree of detail you can see. If you have a larger diameter telescope, you can collect more radiation, and you can also see tinier detail. However, the wavelength of the radiation you are detecting also affects the degree of detail achievable. The longer the wavelength, the bigger the telescope has to be to reach the same level of detail. To get the same detail at millimeter wavelengths as you do with an optical telescope, you would have to have a telescope the size of a football field—obviously out of the question. The solution is to use an array of interconnected smaller telescopes. Then the degree of detail depends not on the size of the individual telescopes in the array but on how far apart you can separate them and still get a coherent signal.

Q: And in addition to being used by the collaborating institutions (California Institute of Technology, University of California at Berkeley, University of Illinois at Urbana-Champaign, and University of Maryland), CARMA is accessible to the broad community of astronomers, correct?

A: We believe this is one of the most important things that we will do. Illinois is going to help us by ensuring we have an easily accessible archive, so scientists who may not be black-belt interferometrists will still be able to pose a problem and get an answer. CARMA needs to be accessible to observers on a variety of levels. (See sidebar for more on the CARMA cyberinfrastructure.)

Q: What's particularly advantageous about the new site in the Inyo Mountains?

A: It's about twice as high as the elevation of the previous arrays [approximately 7,500 feet above sea level] so that the deleterious effect of water vapor in the Earth's atmosphere decreases by a factor of two. We also can separate the telescopes by rather large distances. The combination of the increased sensitivity, lower atmospheric opacity, and the increased collecting power of the telescopes, coupled with new or upgraded electronics, means that we can make detailed, wide area maps faster than has been possible with any other instrument to date.

Q: How will CARMA continue to develop?

A: We see ourselves as a place where students and post-docs will come and get firsthand experience with a state-of-the-art instrument. This is really where we expect to teach and inspire instrumentalists and scientists for the future. The CARMA we're building today needs to continue to evolve to stay alive and attract the best researchers of the next generation.

For more information: http://www.mmarray.org/intro.html

CARMA's cyberinfrastructure

Athol Kemball, leader of the radio astronomy imaging team at NCSA, explains the tools to transfer, process, and store data that are an integral part of the CARMA project.

Q: Can you speak for a minute about the need in astronomy in general, and with the CARMA project in particular, for cyberinfrastructure support?

A: Modern telescopes are faced by two new challenges: increasing data volumes and the need to make telescopes easily accessible to the broadest user community. The data rates are being driven exponentially by advances in electronics and detectors. These data rates and the need for open access make custom interactive reduction untenable as a routine means of analyzing the data; smarter scientific workflows are needed. Automated pipeline reduction is essential for new telescopes such as CARMA so that the telescope can be used by the broadest possible astronomical community. Cyberenvironments in turn offer the best means of providing access by the community to the underlying compute and data resources needed for CARMA.

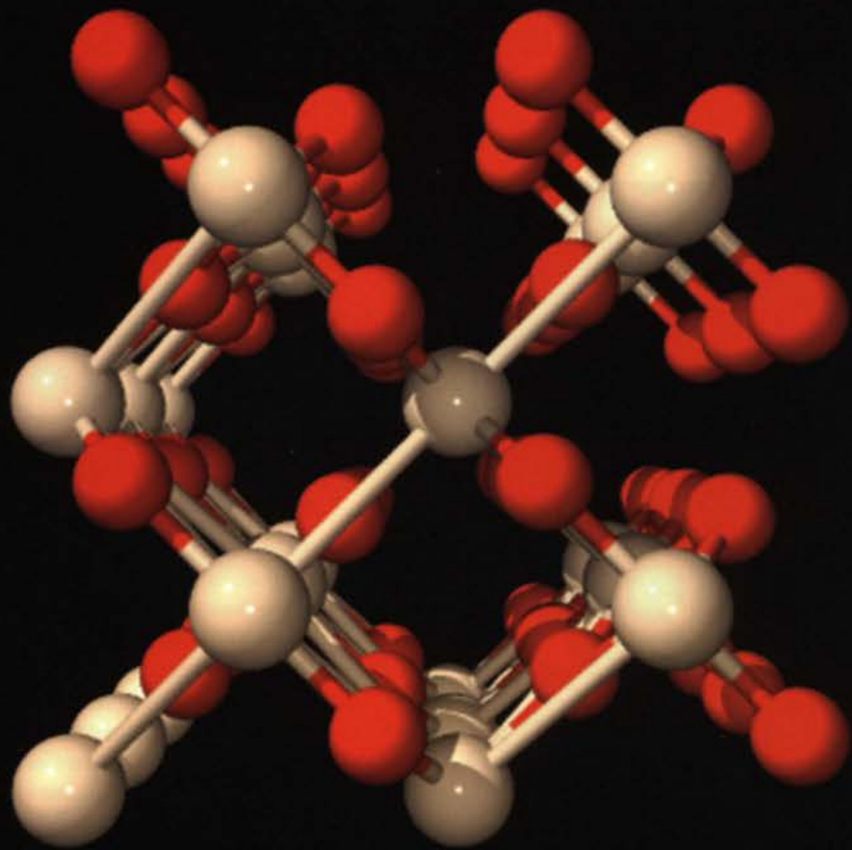
Q: Can you describe the capabilities that were developed for CARMA?

A: The CARMA data archive resides at NCSA, and we will also host the CARMA data reduction pipeline. NCSA has participated closely in the community consortium that built the software system for CARMA, and we lead the team responsible for cyberenvironments, archiving, calibration, and imaging in the CARMA consortium. During construction we contributed to the overall flow of science data and metadata from the telescope to the archive at NCSA and all cyberinfrastructure components involved in this data flow. NCSA also invests in community codes for high-performance calibration and imaging. This application cyberinfrastructure is a vital part of building the automated science workflows for CARMA. Our current focus is on deploying a cyberenvironment to host these capabilities.

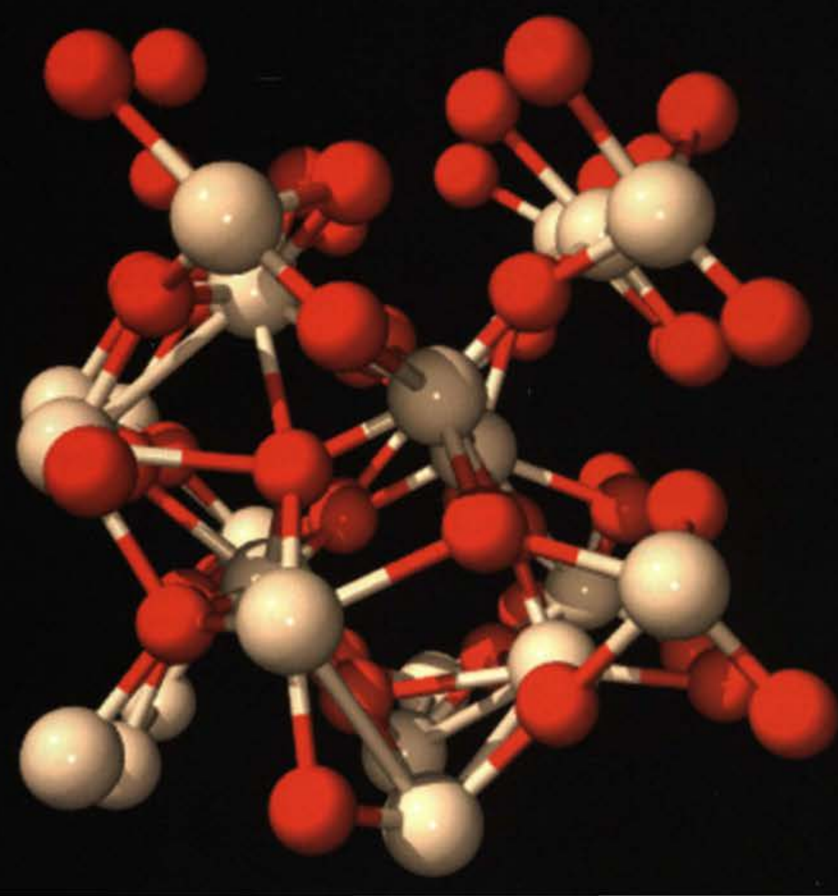
Q: Can you describe the process of developing the CI component of CARMA?

A: The CARMA computing team is a distributed community collaboration, including members from the University of Maryland, University of California at Berkeley, Caltech, the University of Chicago, and NCSA, and it is governed by the requirements of the user community.

In our cyberinfrastructure development we made extensive use of past experience with the archive for the Berkeley-Illinois-Maryland Array (BIMA), which is also hosted at NCSA. We reused infrastructure from this prior project, augmenting this with more recent developments in shared cyberinfrastructure development at NCSA and elsewhere in the cyberinfrastructure community. The success of a project such as CARMA requires extensive reuse of common cyberinfrastructure components wherever possible.



You don’t understand the pressure



by J. William Bell

Geophysicists at the Carnegie Institution of Washington take a crack at the chemistry of minerals deep in the Earth's mantle.

In the late 1950s and early 1960s, "Project Mohole" attempted to retrieve a sample of mantle by drilling a hole through the Earth's crust at the bottom of the ocean. Run by an informal band of scientists called the American Miscellaneous Society and the National Science Foundation, the project returned important information on the crust, but it was discontinued before it breached the Mohorovicic Discontinuity that marks the boundary between crust and mantle.

Project Mohole made it to a depth of about 180 meters, and the record stood at about 2,100 meters in 2005. A Japanese scientific vessel, the Chikyu, is currently being outfitted to close the final 800 meters to the goal.

With 50 years of attempts, it's clear that understanding the earth's mantle is a challenge that will never be met by direct observation alone (or even direct observation primarily). Once the Mohorovicic Discontinuity is tapped, researchers will still only kiss the outer realms of mantle that continues to a depth of 2,900 kilometers. Instead, geophysicists rely on indirect means.

Razvan Caracas and Ronald Cohen of the Carnegie Institution of Washington in Washington, D.C., plumb the mantle's greatest depths by predicting properties of minerals and melts at high pressures and temperatures using NCSA's Cobalt and Copper supercomputers. They perform first-principles computations within density functional theory. These calculations are called "first-principles" because there is no experimental input—they use only fundamental physics as described by quantum mechanics and electrodynamics in the form of the density functional theory. Properties of Earth's materials are obtained from simulating the behavior of electrons and nuclei under certain thermodynamic conditions.

"We can't go out and look at samples," says Cohen. "We only have seismological studies and experiments and simulation. We have no other information on how the deep Earth might work."

With their studies, they provide key data about the physical properties of minerals under extreme conditions of temperature and pressure. These data are further used by other geoscientists to interpret observations, to plan experiments, or to build geodynamical models of the interior of the planets.

"The properties that can be seen [by experimentalists or by those doing computer simulations] are limited. We try always to work closely with one another and put constraints on one another...Sometimes [a feature] is first predicted by experiment. Other times, in theory. But the constraints give us an idea of what to look for, and one method is often an immediate help to the other," says Caracas, a Carnegie Fellow at the institution who runs the simulations.

Comparison between the crystal structure of silica. The top panel is the structure under static conditions (0 Kelvin and 120 gigapascals). The bottom panel is a snapshot of the molecular dynamics run showing the dynamic structure of the same material under thermodynamic conditions similar to the Earth's deep mantle.

A puzzle at 20 million psi

Caracas and Cohen focus not on the upper boundary of the mantle but on the opposite side: the core-mantle boundary. This boundary between the solid, silicate-based mantle and the liquid, iron-based core is "the most dramatic interface in the whole Earth," according to Caracas. The contrast in chemistry, mineralogy, viscosity, and seismic wave velocities is greater even than the contrast between ocean water and the solid bottom. Pressure in this region is more than 130 gigapascals, or about 20 million pounds per square inch. Temperature is more than 2,500 Kelvin.

The mantle side of the core-mantle boundary, the D" (dee double prime) region has "puzzled geophysicists for decades," according to a 2005 article in *Geophysical Research Letters* by Caracas and Cohen. As with other boundary areas in the earth's mantle, seismic waves behave differently when they cross it. The why of this puzzle lies in the peculiarities and the complexity of the signals received from this region and in its remoteness and extreme nature of its thermodynamic conditions.

For years, it was thought that the bulk of this lowermost mantle was composed of an iron-bearing magnesium silicate with perovskite structure. Perovskite ($MgSiO_3$) is not found on the surface, but it makes up about 70 percent of the mantle. However, diamond-anvil cell experiments, which compress a sample between two diamonds to recreate the extreme pressures inside planets, complicated the conventional wisdom in 2003. They showed that the silicate undergoes a phase transition at about 125 gigapascals and 2,500 Kelvin. The crystal structure of perovskite is altered, and it transforms to what is called post-perovskite.

This physical change offers possible explanations for the mystifying behaviors ascribed to the lowermost mantle; seismic waves traveling through crystal structures in the post-perovskite formation would have a different signature than waves traveling through perovskite, for example. But without a clearer understanding of what influences the change and what impacts it has on the mantle's character, those explanations remain broad and subject to further change. That's where Caracas, Cohen, and their collaborators come in.

Pick up the change

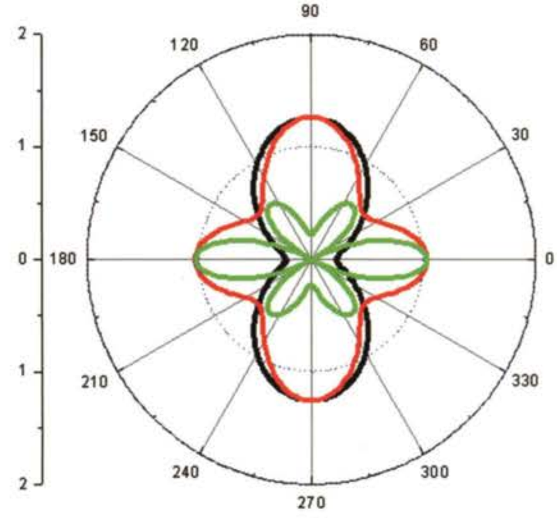
One of the prime targets of their simulations is the effect that the addition of certain atoms has on perovskite's transition into post-perovskite. The team has shown that aluminum oxide, which is known to be present in the lowermost mantle, slightly increases the pressure at which the transition takes place. The introduction of additional iron ions, meanwhile, considerably reduces the transition pressure. These results were published in *Geophysical Research Letters* and presented at the Annual Meeting of the American Geophysical Union in Fall 2005.

More thorough calculations of these changes, now underway at NCSA, will allow the team to build complete phase diagrams of the perovskite to post-perovskite transition and to assess how iron influences that transition. A phase diagram shows the point at which the transition occurs at different temperatures and pressures. At a higher temperature, the required pressure might be lower, for example.

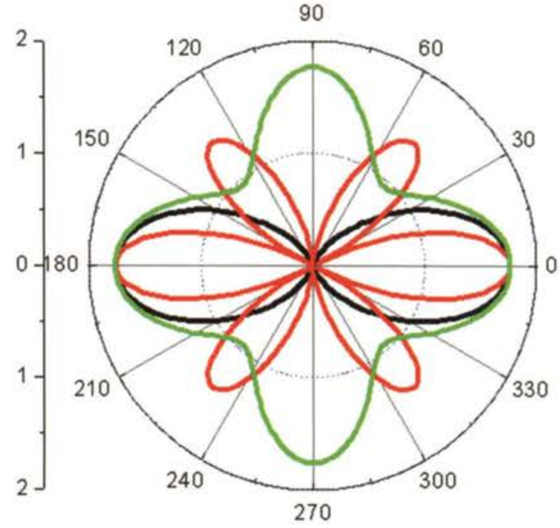
The team is also calculating the elastic constants of the crystals. This feature equates to the stiffness of the crystals as they deform and return to their original shape under the sort of stress that might be generated by a seismic wave.

In 2006, meanwhile, the team published on the Raman spectra of perovskite and post-perovskite—the first time these spectra had ever been derived for a mineral by simulation using density functional theory. Raman spectroscopy is an experimental technique used to study the vibrational characteristics of materials. When light hits a molecule or a crystal, the incoming photons interact with

Absolute seismic anisotropy of the $MgSiO_3$ and $FeSiO_3$ perovskite and post-perovskite as computed at 120 GPa. (Absolute seismic anisotropy is the absolute difference between the velocities of shear waves propagating in the vertical and in the horizontal planes.) Post-perovskite has a lager seismic anisotropy and its presence in the D" layer may explain similar features recorded from that part of our planet.



$MgSiO_3$ perovskite structure



$MgSiO_3$ post-perovskite structure

particular atomic vibrations and change their original frequency. This shift is a direct expression of the atomic vibrational frequency. It is measurable by Raman spectroscopy and is very specific. Using it, scientists can determine the interatomic chemical bonds that are present in the molecule or crystal and can use it as a nondestructive identification tool.

The team's article, which also appeared in *Geophysical Research Letters*, showed substantial differences between the spectra exhibited by perovskite and those for post-perovskite. This finding tells experimentalists that Raman spectroscopy is a good means of testing whether the transition has taken place in a sample, because the disparity between the two spectra is so pronounced.

"NCSA's entire approach was essential to delivering these findings," says Caracas. "They helped with setup and compiling our codes, and their flexibility in using the machines and giving us dedicated time was key."

This research is supported by the National Science Foundation's Division of Earth Sciences.

Access Online: http://access.ncsa.uiuc.edu/Stories/core-mantle

For further information: http://www.gl.ciw.edu/

http://www.gl.ciw.edu/~r.caracas/

Team members

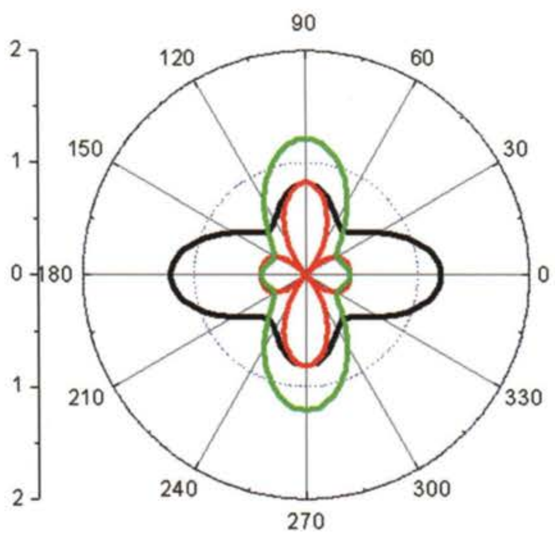
Ronald Cohen
Razvan Caracas
Burkhard Militzer

Using the system that suits you

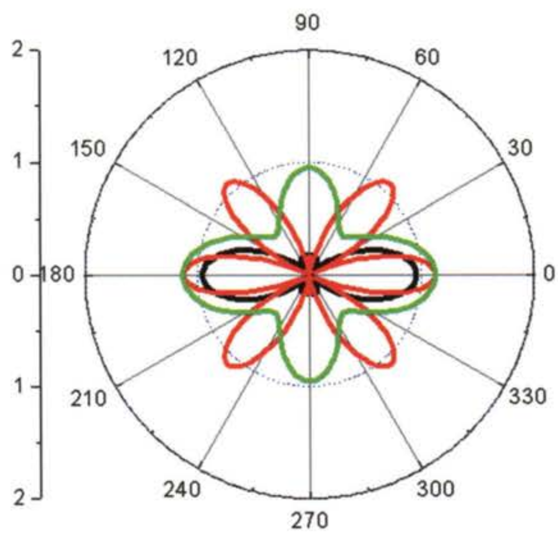
At the incredible pressures like those in the Earth's mantle, some materials undergo a change in their crystal structure. Perovskite becomes post-perovskite around that 120 gigapascal mark, but it stays solid. Other materials, meanwhile, are going through phase changes that we're more familiar with here on the surface. They melt, shifting from solid to liquid just like an ice cube clenched in your fist. These changes influence the nature of mantle and its geophysical properties. So it's little surprise that the team from the Carnegie Institution is interested in those changes, too.

They started with the study of liquid silica (SiO_2). Once it has melted, silica is a particularly challenging area of focus. In fact, it marks the first time the team has done work from first principles on a liquid. "You don't have the regular periodicity of the structure [that you do in a crystal's solid state]," explains Cohen. "Something as simple as the equation of state, which is just the density of the material as a function of pressure and temperature, is much more difficult in many ways. For a crystal you could get that experimentally, with x-ray diffraction techniques. But a liquid doesn't have sharp diffraction, so it's much more difficult."

Simulations of the conditions under which silica melts are conducted on NCSA's Copper and Cobalt systems. One code looks at silica in the liquid phase, while another looks at it in the solid phase. "ABINIT [which is used for the solid] is good for Cobalt [an SGI Altix shared-memory system]. PWSCF is most appropriate for Copper, because the code has already been well adapted to this IBM system. It's very useful to have both styles available at one place," says Caracas.



$FeSiO_3$ perovskite structure



$FeSiO_3$ post-perovskite structure

TECTONIC SHIFTS

by Kathleen Ricker

A new earthquake engineering cyberenvironment could change the way communities prepare for the worst—and help translate science into practical results more quickly.

In the winter of 1811–1812, the New Madrid seismic zone—a rift in the earth's crust that stretches across parts of Illinois, Missouri, Arkansas, Tennessee, and Kentucky—generated a series of earthquakes now estimated to have been between 7.8 and 8.2 on the Richter scale. The massive temblors reversed the flow of the Mississippi River for several days, created lakes in Missouri and Tennessee, shook the White House in Washington, D.C., and rang church bells in Boston, Massachusetts.

Almost two centuries later, the New Madrid seismic zone contains large population centers that are crucial economic and industrial hubs. A major earthquake could result in tens of thousands of human casualties, with thousands dead in southern Illinois alone. Economic loss, both regionally and nationally, could be anywhere between $100-200 billion. And the population displacement caused by a temblor of the size of the 1811–1812 quakes could, by some estimates, be greater than that caused by Hurricane Katrina.

"The amount of energy released by these earthquakes is beyond normal imagination," says Amr Elnashai, director of the Mid-America Earthquake Center at the University of Illinois at Urbana-Champaign, where research is focused mainly on the New Madrid seismic zone and its hazard. "We would be looking at roughly 10 times the area affected by a similar magnitude earthquake on the West Coast." Metropolitan centers like Memphis and St. Louis would be severely damaged, as would the surrounding counties in Illinois, Kentucky, Arkansas, Ohio, and Indiana; even Chicago high-rises would be affected. Closer to ground zero, everything would be severely damaged, including electrical grids, water purification plants, nuclear reactors, and sewage systems.

MAEviz visualization showing damage to hospitals, schools, fire stations, and other buildings and bridges in Memphis, Tennessee, after a hypothetical earthquake. The vertical bars indicate the amount of damage to the structure, with red bars indicating probability of collapse and white bars indicating probability of no damage. Information about the MLGW gas network repair rate at various locations is also included. Users may superimpose as many layers of detail as desired.

Buildings damaged in the Kocaeli earthquake of August 17, 1999. Measured at 7.4 on the Richter scale, it is the most powerful earthquake ever to hit Turkey, killing more than 15,000 and incurring more than $5 billion U.S. dollars worth of building loss. Working with the city of Istanbul, NCSA and the MAE Center are currently creating a Turkish version of MAEviz to help the city prepare for future seismic events.

Worst-case scenario

The New Madrid seismic zone has more earthquake activity than any other region east of the Rocky Mountains, although the quakes are minor and cause little damage. However, federal, state, and local governments, as well as utility companies and businesses, are responding to the possibility of another major quake by retrofitting bridges, buildings, pipelines, railways, highways, and other crucial infrastructure such as nuclear power plants and electrical power stations. But with continual urban growth and redevelopment, deciding where best to apply limited resources is difficult, especially when taking emergency response into account. Preparation and crisis management require access to the latest science and other kinds of information that can change rapidly and frequently. Moreover, because seismologists are still exploring the unique characteristics of the New Madrid fault itself, such as the ways in which the thick Mississippi River sediments amplify and liquefy ground motion, new findings can have a major impact on disaster planning.

To provide a conduit between researchers and practitioners, the Mid-America Earthquake Center (MAE) and NCSA are jointly developing the MAEviz loss assessment system, a cyberenvironment that gives decision makers access to tools aimed at helping them assess risk and determine how to allocate resources for mitigating risk. Based on technology developed at NCSA and the University of Michigan, MAEviz has a modular structure and includes analytical tools and data repositories for modeling earthquake damage. One module simulates the hazards posed by earthquakes of different intensities and durations. Others provide detailed inventories of bridges, pipelines, networks, railways, streets and highways, and buildings, including essential facilities like hospitals and schools, which allow users to assess potential loss. Combining these modules into a workflow with a third type that encodes knowledge about fragility (the probability of damage to facilities and networks) produces loss estimates that still other modules translate into terms of human and economic loss, demonstrating where and how reducing fragility can best reduce those losses, and helping decision makers determine the most cost-effective retrofits.

MAEviz's modular architecture not only makes it more flexible than other loss assessment software, but also more dynamic. As researchers produce algorithms and data that can provide more accurate estimates, the new modules are swapped in so decisions can always be based on the best science available. And because MAEviz is a network-aware environment, new software and data can be downloaded and new results uploaded for sharing between collaborators through a secure web portal. If disaster strikes, plans can be retrieved from the network and damage assessments can be recomputed using the latest information.

Putting science into practice

For Jim Myers, chief architect of MAEviz and leader of NCSA's Cyberenvironments and Technologies Directorate, MAEviz is an example of how cyberenvironments can accelerate the interaction between scientific research and its application in the field. "Five years ago you would have built a MAEviz as an application, and you would have just accepted that the scientific knowledge behind it would grow increasingly outdated," Myers says. "While people might have heard about better methods or newer data at conferences and read about them in publications, they wouldn't have been able to do more than idly speculate about how the new information might affect planning. Why shouldn't you be able to just plug in a new module and immediately see the potential impact of the latest research?"

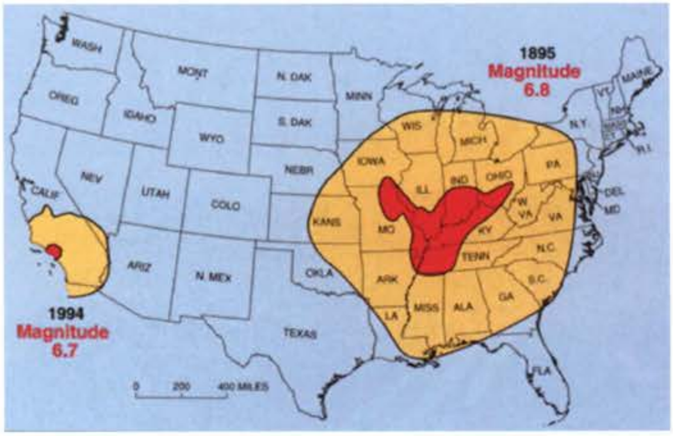
Myers and the MAE researchers envision MAEviz as a collaborative bridge between researchers, engineers, and planners. Researchers create new analysis modules and data sets, engineers create and post scientifically rigorous scenarios based on the new information, and planners download appropriate scenarios to assess the impact on strategy. If the potential impact is significant, all three groups can coordinate to speed additional research and validation efforts, resulting in informed decisions.

Recently, MAE conducted a pilot collaboration with a regional utility, Memphis Light, Gas, and Water (MLGW). A customized version of the software, MAEviz-MLGWgas, was used to assess potential earthquake damage to the company's existing cast-iron piping network—a small portion of their gas network, but substantial enough to demonstrate MAEviz's potential as a loss assessment tool and an asset management environment for MLGW. Since then, MAE has also collaborated with other partners to provide modules customized for different areas and needs, including the transportation network for Charleston, South Carolina. MAE is currently working with FEMA to create a comprehensive loss assessment model in the event of a New Madrid earthquake. That model uses MAEviz in combination with FEMA's broader, more general loss assessment tool, HAZUS.

MAEviz is attracting interest from other parts of the world as well, especially earthquake-ravaged areas such as Turkey, Pakistan, and Indonesia. The MAE Center is currently partnering with Istanbul Technical University and the Municipality of Istanbul to build a similar framework with a Turkish-language interface for that city. A similar project is underway for the city of Islamabad, Pakistan, under the direction of Jerome Hajjar and Arif Masud, both at MAE, and funded by the Pakistani government's Earthquake Reconstruction and Recovery Agency and conducted jointly with the National University of Science and Technology in Islamabad. Its goal is to estimate losses from earthquakes like the one in October 2005 that killed almost 75,000 people and left 2.8 million homeless in Pakistan.

"MAEviz is enabling collaboration between seismologists, social scientists, structural engineers, technical engineers, loss assessors, and emergency managers," says Elnashai. "It really opens things up in a big way."

The development of MAEviz is funded by the National Science Foundation through support of the Mid-America Earthquake Center, an NSF Research Center.



Map depicting damage sustained by two earthquakes of similar magnitude: the 1895 Charleston, Missouri, earthquake and the 1994 Northridge, California, earthquake. While fewer earthquakes occur in the Midwest than on the West Coast, the extent of damage is much greater. Red indicates damage to buildings; yellow indicates areas where shaking could be felt. *Image courtesy of the US Geological Survey.*

Access Online: http://access.ncsa.uiuc.edu/Stories/MAEviz

For further information: http://maeviz.cee.uiuc.edu/

http://mae.ce.uiuc.edu/software_and_tools/maeviz.html

Team members

Amr Elnashai
Tom Finholt
Terry Fleury
Jerry Hajjar
Shawn Hampton
Jong Sung Lee
Terry McLaren
Jim Myers
Chris Navarro
Greg Peters
Bill Spencer
Nathan Tolbert

The MAEviz portal and web application depend upon a variety of open source software developed at NCSA, Pacific Northwest National Laboratory, and the University of Michigan as well as the expertise and support of their development teams.

Data to Knowledge (D2K)
http://alg.ncsa.uiuc.edu

Scientific Annotation Middleware (SAM)
http://collaboratory.emsl.pnl.gov

Tupelo
http://dlt.ncsa.uiuc.edu

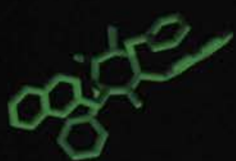
MyProxy
http://grid.ncsa.uiuc.edu/myproxy

Sakai
http://www.sakaiproject.org

Visualization Toolkit (VTK)
http://www.vtk.org

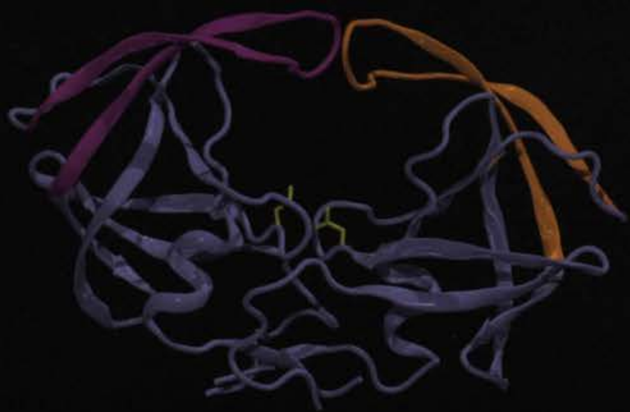
The quest for a cure

Story by Barbara Jewett
Infographic by Blake Harvey



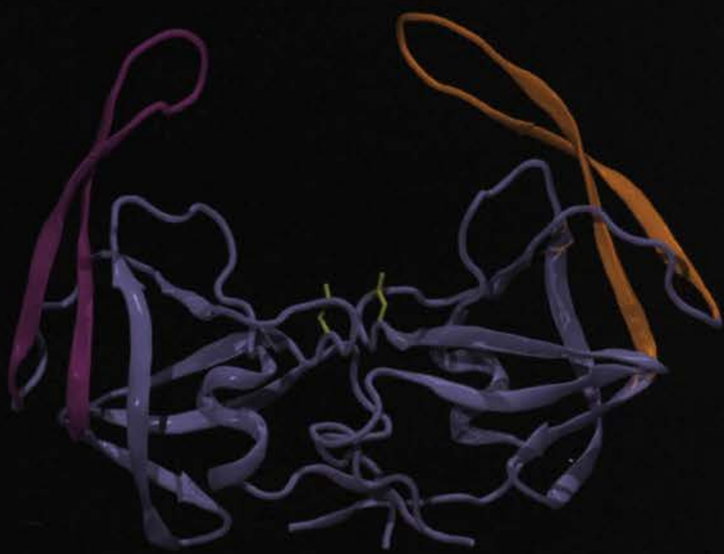
Due to its central role in processing viral polypeptide precursors, HIV-1 protease (HIV-PR) continues to be one of the primary targets of anti-AIDS drug discovery. A greater understanding of the mechanistic events associated with binding of HIV-PR substrates and inhibitors is critical for the design of more potent inhibitors of the enzyme.

Carlos Simmerling and his colleagues at Stony Brook University are viewing the moments at which "starter molecules" for HIV are most vulnerable to new drugs. Using NCSA's SGI Altix supercomputer, called Cobalt, they successfully simulated how HIV protease changes between two forms that already have been determined through experiments. More importantly, however, the group captured the protease in a third, fully open state—one that had previously been hypothesized but never directly observed.



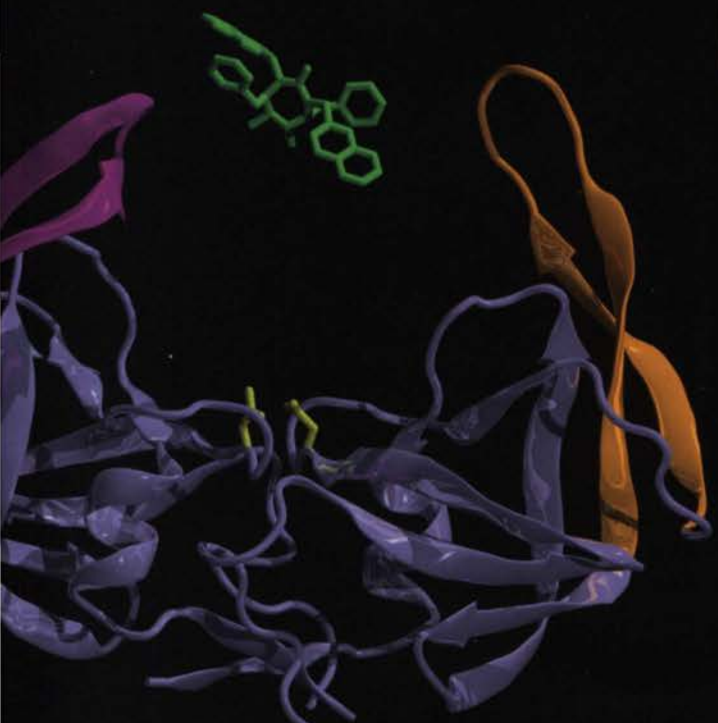
1 Coming together

HIV protease is a dimer, formed when two identical chains come together (right and left). The semi-open form is shown, as seen in the crystal structures without an inhibitor. The two flaps (purple and orange, one from each monomer) contact each other and cover the binding pocket, blocking access to the active site. The aspartic acid side chains that the protease uses to carry out its function are at the bottom of the binding pocket (shown in yellow). During simulation without an inhibitor, the flaps are very flexible and the flap tips periodically lose contact with each other.



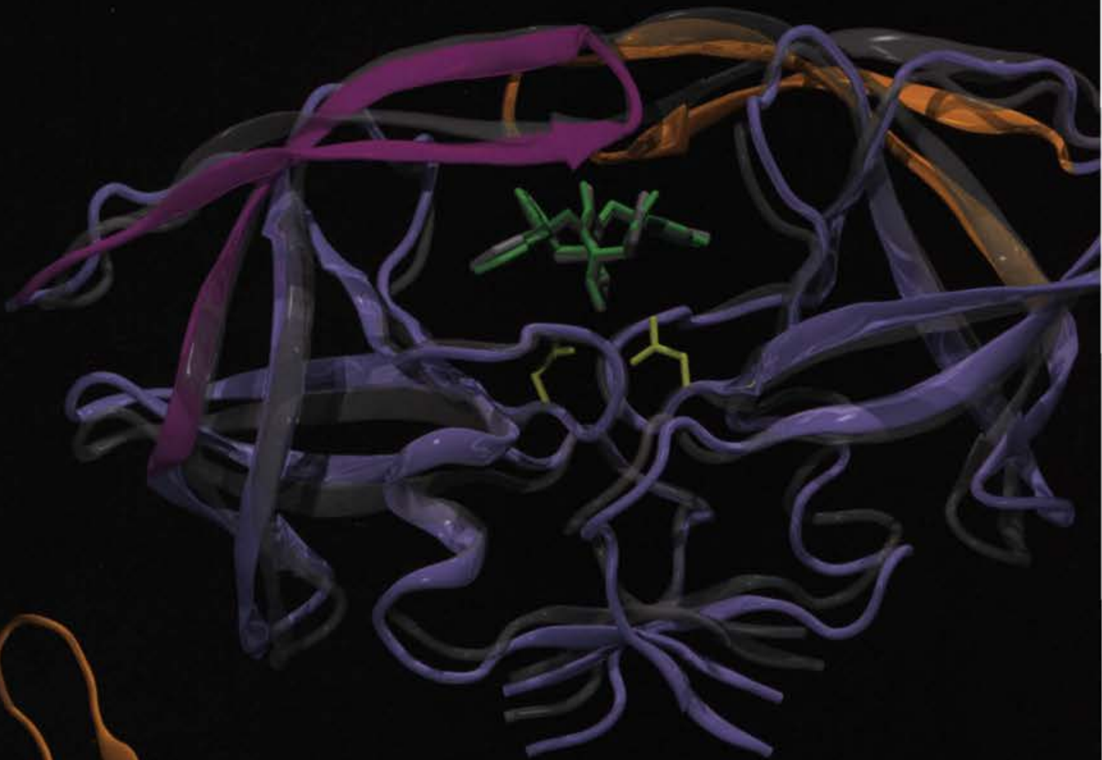
2 Flipping flaps

After the flap tips separate, the two monomers rotate with respect to each other and the flaps become separated by about 20 angstrom, opening the active site and allowing substrate or inhibitors to enter. After a short time, the structure closes back to the semi-open form in which it spends most of its time. The binding pocket is again inaccessible to substrate or inhibitors. Occasionally the flap tips swap sides—the left flap (purple) is in front, while in the normal semi-open form the right flap (orange) is in front. Having the left flap in front is observed in crystal structures of the inhibitor-bound protease, but without an inhibitor this conformation is unstable and rapidly switches back. NMR experiments had suggested that the protease might undergo this kind of transient flap reversal.



3 In the pocket

In the open form, substrates or inhibitors can enter the binding pocket. In this simulation, an inhibitor (green) was manually placed near the entry to the active site of an open structure from a simulation without an inhibitor. The inhibitor enters the binding pocket and the flaps begin to close. The inhibitor forms specific hydrogen bonds with one of the catalytic aspartic acids and one flap, accelerating the closing process. The other flap closes and helps to pack the bulky side groups of the inhibitor into the pocket.



4 A new stability

The protease becomes fully closed. For comparison, the crystal structure of the complex is shown as a transparent gray model. The simulation model is nearly identical to the crystal structure even though the closed form was not used for the simulations. Once the protease closes with an inhibitor in the pocket, the flaps become stable and no longer open in the simulations. Note that the left flap (purple) is in front when the inhibitor is bound. When the protease closes without an inhibitor, the right flap is in front. This flap reversal upon inhibitor binding is also observed in the crystal structures.

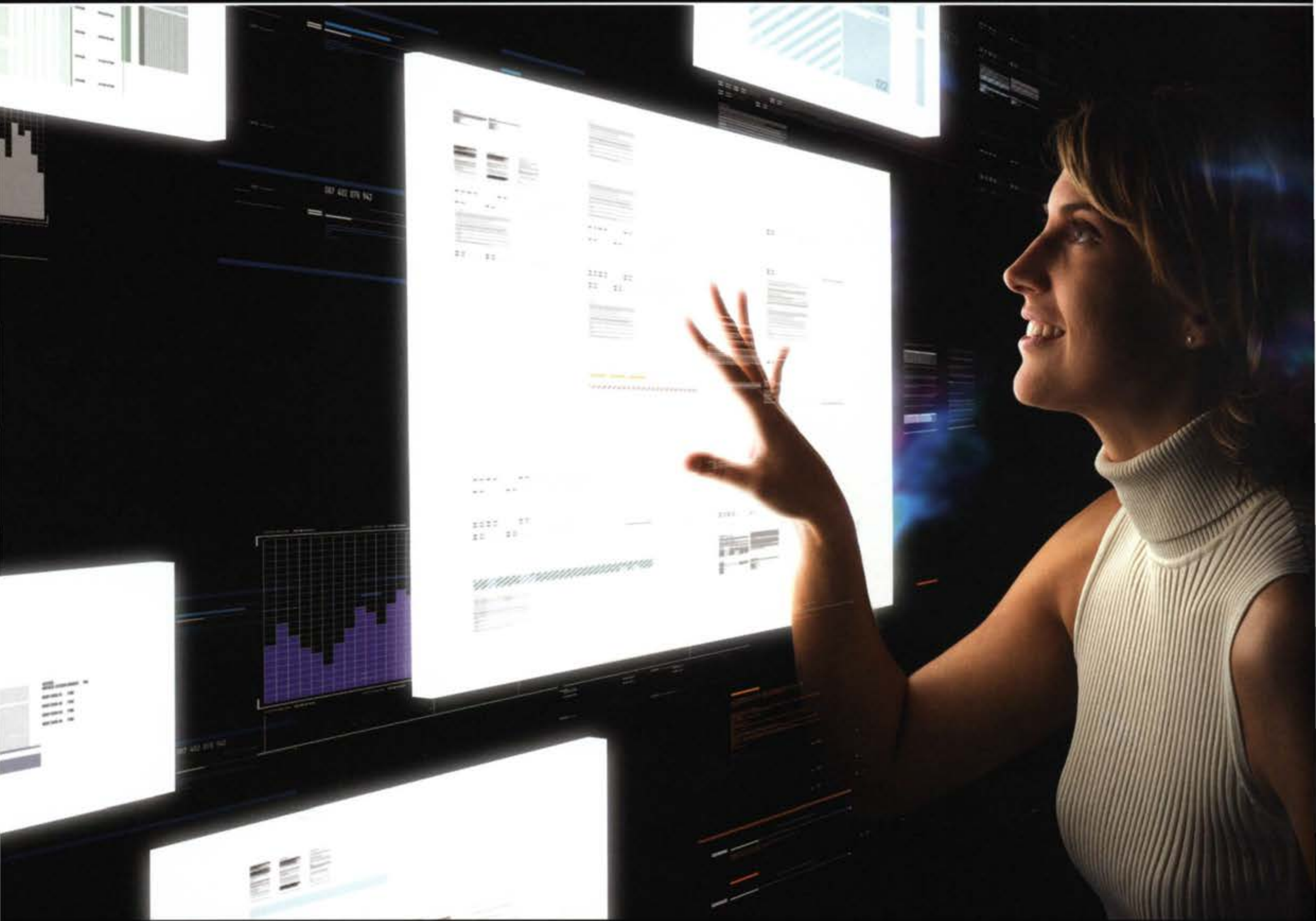
The potential for overcoming HIV

The change in flap handedness upon ligand binding is clearly reproduced in the team's simulations, but its possible role in HIV-PR function remains unexplained. Understanding the issues that govern HIV-PR flap mobility has profound implications for elucidating the detailed mechanism of this enzyme and in the design of new therapeutic agents, such as allosteric inhibitors intended to interfere with flap opening and thereby with enzymatic function. The team's work was published in the *Proceedings of the National Academy of Sciences* and in the *Journal of the American Chemical Society.*

This work was supported by the National Institutes of Health and the Department of Energy.

All images were made using VMD and POV-Ray.

Team members
Viktor Hornak
Asim Okur
Robert C. Rizzo
Carlos Simmerling



Accepting a new set of challenges

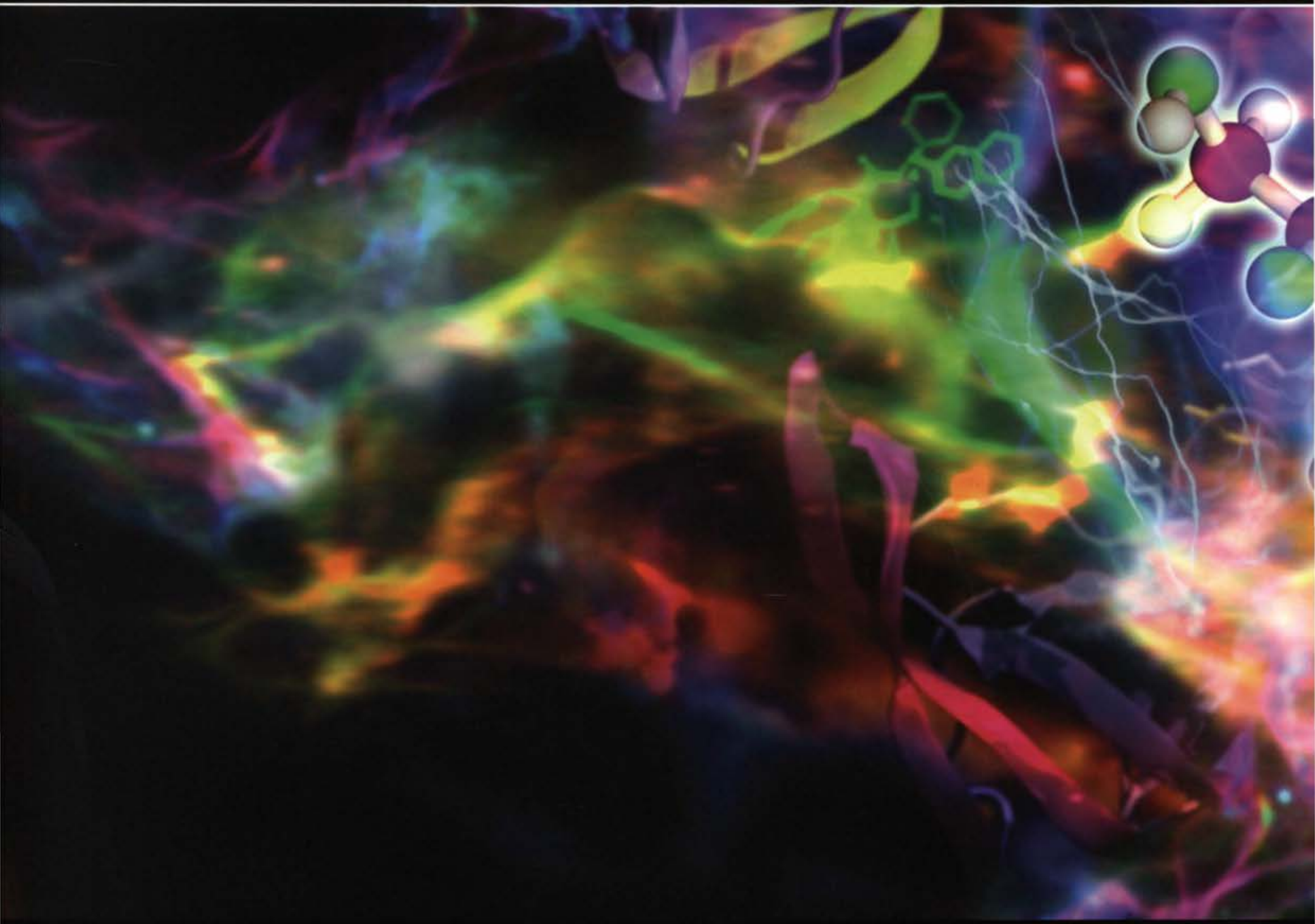
The University of Illinois at Urbana-Champaign creates the Institute for Advanced Computing Applications and Technologies with NCSA.

Computing and information technologies are among the most significant achievements of the 20th century. These advances exert a significant impact on scholarly disciplines, not only in those areas of science and engineering that have provided the initial seedbed but in new and rapidly expanding communities—life sciences, environmental research, business, healthcare, and the social sciences, arts, and humanities.

Far-reaching outcomes for the 21st century will change the way we discover new knowledge, predict the behavior of complex natural and engineered systems, and manage and analyze data. These changes will improve our health, security, and economic welfare. The transformation of data to information to knowledge to applications that benefit society represents a grand challenge.

To realize the full potential of this rapid growth, the University of Illinois at Urbana-Champaign is creating the Institute for Advanced Computing Applications and Technologies (IACAT). The institute will combine faculty-initiated research in its academic units with the advanced technology capabilities at the National Center for Supercomputing Applications (NCSA).

"IACAT will promote synergies between research faculty in different disciplines and NCSA staff," says Thom Dunning, director



of IACAT and NCSA, "positioning both to solve the nation's most challenging problems, provide a supporting cyberinfrastructure that will enable extraordinary research advances, and extend the impact of that research. It will deploy those solutions beyond their original points of inquiry."

By forming alliances with other campus institutes and groups, IACAT will create infrastructure that promises advances that go far beyond the knowledge of any particular researcher or community.

NCSA Inside

The great promise offered by computing and information technologies will best be realized if the experts creating new computing and information technologies and infrastructure are closely coupled with discipline-based researchers who are deeply involved in developing research applications that make full use of these new capabilities. Collaborative multidisciplinary research is at the heart of this activity.

"Multidisciplinary research and development has always been a core strength of our campus, and it is becoming increasingly important as science and engineering tackle the complex problems confronting our nation and our world," says University of Illinois at Urbana-Champaign Chancellor Richard Herman. "The Institute for Advanced Computing Applications and Technologies will combine NCSA's world-class development and deployment of advanced computing technologies with faculty at the University of Illinois who are pushing the boundaries of their applications. It is an incredibly powerful mixture that will profoundly affect the future of research and education."

The unique role of the institute will be to foster deep partnerships with NCSA, which has contributed significantly to the birth and growth of the worldwide cyberinfrastructure. NCSA operates some of the world's most powerful supercomputers and develops the software infrastructure needed to use these systems efficiently. The institute will harness the national cyberinfrastructure and, at the same time, inform the creators of that infrastructure about important application areas where new and improved tools and methods are needed. IACAT, with NCSA inside, will provide unparalleled opportunities to advance domain-specific applications of advanced computing and information technologies while spurring the development of new computing and information technologies.

Lincoln comes to Illinois

Many of the systems at NCSA are supported by the National Science Foundation. They're used by the national research community and allocated by a centralized board that parcels out time on NSF-funded systems around the country. The newest addition to the machine room floor breaks that mold.

Lincoln, an 11-teraflop cluster of Dell PowerEdge™ blade servers, is funded by the state of Illinois. Time on the cluster will be allocated at the discretion of the leadership of the new IACAT. It will be devoted in large part to strategic institute, campus, and state initiatives and will be used by NCSA's private sector partners. This set up allows for additional flexibility in developing new research themes for the institute, deeper relationships with users of computing time, and a more nimble organization.

"The University of Illinois at Urbana-Champaign has identified several major new initiatives this past year, including initiatives in sustainable energy and the environment, integrated sciences for health, and informatics and intends to be a major player in the development and use of petascale computing. All of these initiatives need advanced computing capabilities," says Thom Dunning, director of IACAT and NCSA. "A focus of the institute will be to provide the resources, expertise, and good ideas needed to move these initiatives forward. Lincoln is a crucial part of that goal."

Lincoln is currently capable of about 11 teraflops, or 11 trillion mathematical calculations per second. In the coming months, the system will be upgraded further for greater capability. At that point, it is expected to be the largest supercomputing resource funded by a university for science, engineering, and humanities, and for use by private sector program partners.

Lincoln follows in the path of two other Dell clusters that NCSA offers to the national research community—Tungsten and T2. Tungsten, which is supported by the National Science Foundation and is allocated through NSF's peer reviewed process, is an extremely popular resource, attracting requests for allocations that far exceed the amount of time available on the system. T2 has been extensively used by the center's private sector partners since it was added to the machine room late in 2004.

Research themes

The scope and complexity of advanced computing applications and technologies require close collaboration across many disciplines and skills. IACAT will be organized around a set of research themes that will address major opportunities and challenges that are too complex for individuals, or even small groups of researchers, to tackle on their own.

"Research themes will be defined by combining bottom-up creativity with the strategic positioning of NCSA and the campus," says Dunning. Research themes will be broadly defined and may include projects that span multiple disciplines. A series of workshops, begun in 2005 and continuing through 2008, is exploring possible research themes.

Because different fields have different approaches to research, newly developing communities may take time to emerge and connect with NCSA. Community-building activities will therefore be important in addition to traditional research. It is expected that research themes and projects will serve to position university and NCSA researchers for emerging opportunities. For example, strategic initiatives recently announced by the university include:

- The Illinois Sustainable Energy and Environment Initiative, which will use the campus' unique interdisciplinary strengths in science and technology, economics, the humanities, and the social sciences to develop new technologies, practices, and policies in the sustainable use of vital resources such as energy, water, and land.
- The Integrated Sciences for Health Initiative, which will apply Illinois' expertise in the physical sciences, engineering, and information, medical, and life sciences to improve human health.
- The Illinois Informatics Initiative, which is an integrated approach to prepare students for the information-technology enabled workforce in the natural sciences, humanities, social sciences, and the arts and on decision support for business and government.

Resources

The institute will be staffed by faculty from university academic units and by NCSA academic professionals. The university is making new resources available to launch the Institute for Advanced Computing Applications and Technologies. The institute will work in close partnership with participating academic units beginning in the 2006-07 academic year to identify and recruit institute faculty. The institute will share the new 140,000-square-foot building with NCSA.

"This university has a great history of support for computing that provides high-end resources and tools that scientists and other researchers need to answer their most challenging questions," says Herman. "The Institute for Advanced Computing Applications and Technologies will target those areas that stand to make the most of that support. It will be a gift to future generations of researchers, ensuring that the work that they do will take advantage of the continuing revolution in computing."

For further information: http://www.iacat.uiuc.edu/

LEAD WEATHERS FIRST BIG TRIAL

by Barbara Jewett

A multi-disciplinary effort involving nine institutions—including NCSA—and more than 100 scientists, students, and technical staff, LEAD comprises a complex array of services, applications, interfaces, and local and remote computing, networking, and storage resources that can be used in a stand-alone fashion or linked together in workflows to study mesoscale weather (meteorological phenomena approximately 2-200 kilometers in horizontal extent, such as thunderstorms and squall lines).

The LEAD (Linked Environments for Atmospheric Discovery) framework provides users with an almost endless set of capabilities ranging from simply accessing data and perhaps visualizing it, to running highly complex and linked data ingest, assimilation and forecast processes in real time and in a manner that adjusts dynamically to inputs as well as outputs. The methodologies and infrastructures being developed are extensible to other fields, including medicine, ecology, hydrology, geology, oceanography, and biology.

"We had dedicated time on both NCSA's Mercury and Tungsten clusters, using 256 processors of each machine" says Jay Alameda, NCSA project manager, of the LEAD trial that took place July 13 as part of a Unidata workshop at the National Center for Atmospheric Research in Boulder.

Following a tutorial, users configured their accounts, developed their forecasts, and launched the associated workflows on the TeraGrid. According to Alameda, two runs were made, with 25-30 users simultaneously in each run. Once both runs were complete, users came together in a plenary session to view and discuss their results.

"Some aspects of the project worked very well; other parts, we discovered, did not," says Alameda. Teams are identifying the modifications that will be needed for those areas.

Participants' overall impression was that the capabilities provided by LEAD were useful. The University of Michigan's Katherine Lawrence and her community engagement team were present to observe the use of the orchestration systems and identify further usability issues related to functional navigation and the use of documentation. The team also recorded explicit suggestions for improving the consistency of software with existing conventions and expectations of meteorologists and students (for example, preferred notation for specifying date and time).

The workshop afforded a unique opportunity to unveil the rapidly maturing capabilities of LEAD to a select group of university researchers and faculty who are particularly interested in running forecast models and applying related tools in the classroom. In addition, the team obtained both quantitative and qualitative feedback and conducted the first scalability and stability tests of the entire LEAD system, including in particular the ability of the TeraGrid to handle the simultaneous submission of dozens of weather research and forecasting jobs.

Participating institutions

Colorado State University

Howard University

Indiana University

Millersville University

University of Alabama in Huntsville

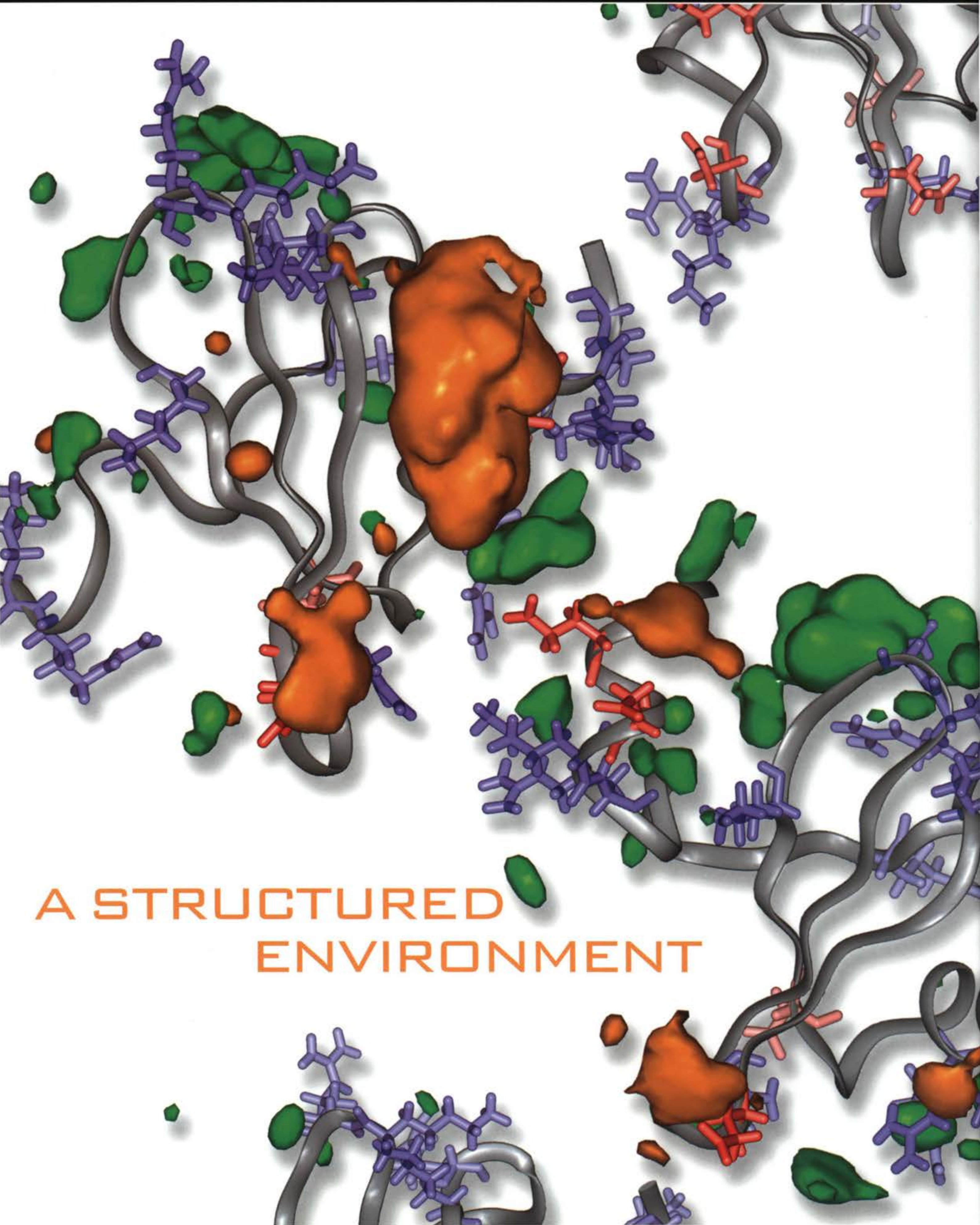
University Corporation for Atmospheric Research (UCAR) Unidata Program

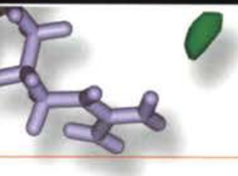
NCSA/University of Illinois at Urbana-Champaign

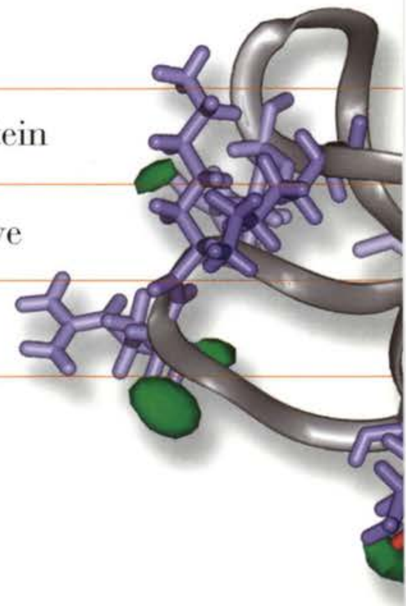
University of Oklahoma

University of North Carolina-Chapel Hill

A STRUCTURED ENVIRONMENT



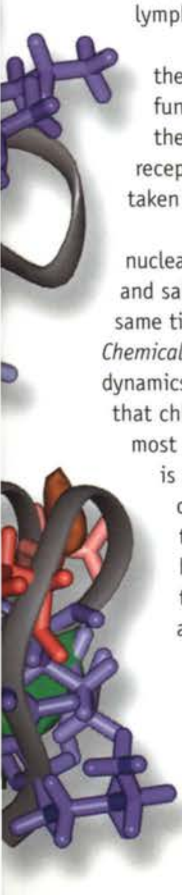




by J. William Bell

Researchers use NCSA's Tungsten to investigate the protein lymphotactin, the gating pathways of mechanosensitive channels, and an allosteric signaling protein.

A team of researchers from the University of Wisconsin at Madison are using NCSA's Tungsten supercomputing system to investigate lymphotactin, a protein manufactured by progenitor T-cells in the human body. Lymphotactin has been linked to autoimmune diseases and to the rejection of transplanted organs. Led by Qiang Cui, an assistant professor in the chemistry department, they are simulating the impact that temperature and salt concentration have on lymphotactin's structure.



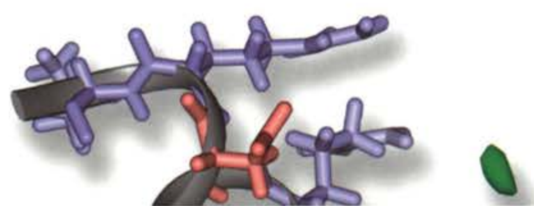
Protein structure can be influenced by such factors in the surrounding solvent, and it heavily influences protein function. Lymphotactin, for example, is thought to stimulate the immune system by binding to specific protein-coupled receptors on immune-system cells. It can only do that if it has taken on specific shapes.

Contemporary experimental studies of the protein using nuclear magnetic resonance imaging tend to change temperature and salt concentration of lymphotactin's environment at the same time. In a paper published in the *Journal of the American Chemical Society* in July 2006, the Wisconsin team's molecular dynamics simulations treated these variables separately. They found that chloride from the salt in the solvent is more attracted to most sections of the protein at higher temperatures. In what is known as the "C-terminal helical region," however, both chloride and sodium distributions are higher at lower temperatures. The team also found that the C-terminal helix partially melts at higher temperatures, regardless of the amount of salt, whereas another region begins to form a helical structure at higher salt concentrations.

"These explicit solvent simulations on the order of 70 nanoseconds would not be possible without the generous support of computational resources from NCSA. The user-friendly infrastructure of NCSA has not only provided the facility for carrying out the molecular simulations required by our projects but also stimulated us to pursue more challenging problems in the biophysical and biomedical area," says Cui.

For example, the Cui group has also been using the NCSA resources to analyze the energetics associated with the conformational transitions in the molecular motor myosin. Specifically, free energy simulations are used to probe the energetic coupling between the conformation of the lever arm and the active site. Although these structural motifs are separated by more than 40 angstrom, their motions have to be tightly coupled for myosin to avoid futile ATP hydrolysis and maintain a high working efficiency. Considering the large size of the system (about 800 amino acids), quantitative free energy simulations are impossible without the NCSA resources. These simulations will lay out a solid foundation for the development of coarse-grained models that capture the long-time behavior of myosin.

Though these simulations only reveal the initial stages of lymphotactin's structural change, they demonstrate the significant impact that environmental variables can have. The team is now in the process of performing more extensive simulations for lymphotactin's other structural form, also under different salt concentrations and temperatures.



"Nearly every single backbone hydrogen bond has to be broken in the structural conversion process," says Cui. "It's truly remarkable that these striking changes—along with the tendency to dimerize, or bond two identical copies of the protein to one another—are facilitated by something as simple as temperature and salt." Although it is unrealistic to expect to observe the entire structural transitions during the molecular dynamics simulations, hundreds of nanosecond simulations are expected to be very instructive regarding how environmental conditions stabilize different structural forms.

The team also recently published NCSA-based simulations of the gating pathways of mechanosensitive channels of large conductance in two bacteria using a finite element method, as well as simulations showing the activation mechanism of a small signaling protein that exhibits allosteric characterisitics, in the June 2006 issue of *Proteins* and the August 2006 issue of *Biophysical Journal*. Large-scale conformational transitions are crucial to the function of many proteins, yet the underlying mechanism remains elusive for most cases. This is because these transitions typically span a broad range of time and spatial scales, and therefore are difficult to characterize for both experiments and computations.

The resources at NCSA are crucial to the development and application of novel computational techniques that can make a major contribution in this important area of biophysics. For example, the reported study of the mechanosensitive channel illustrated the great potential of continuum mechanics in describing functional transitions in proteins in response to the external perturbations like membrane deformation. At the same time, it also highlighted that a key challenge is to develop strategies to smoothly connect atomistic simulations with continuum mechanics models, which is an active area of research in the Cui group as well as for his collaborators. In fact, says Cui, the work in this area done by Xi Chen and colleagues in the Civil Engineering and Engineering Mechanics Department at Columbia University was crucial to the mechanosensitive channel simulation.

This research is supported by the National Institutes of Health, the National Science Foundation, the Research Corporation, and the Alfred P. Sloan Foundation.

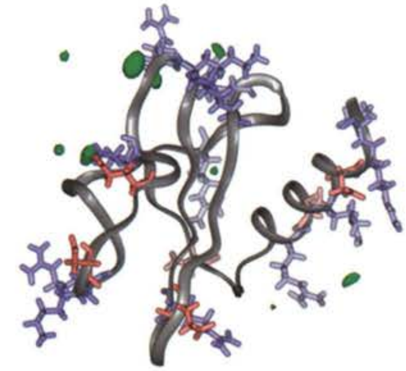
Access Online: http://access.ncsa.uiuc.edu/Stories/lymphotactin

Team Members

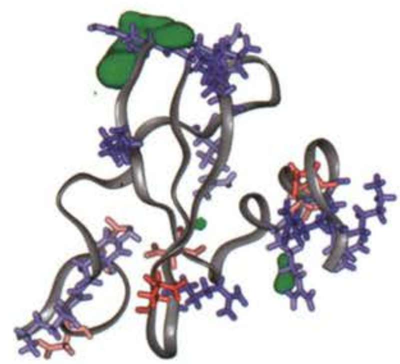
University of Wisconsin at Madison
Quang Cui
Mark Formaneck
Liang Ma
Yang Yang
Jejoong Yoo
Haibo Yu

Columbia University
Xi Chen

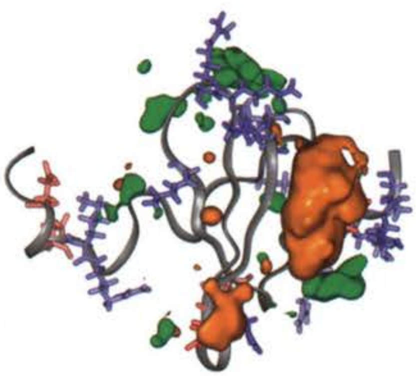
Three-dimensional ion distribution averaged over the entire course of the explicit solvent-ion simulations of lymphotactin, pictured with the final snapshot for each simulation condition. ("n" and "s" indicate no salt and 200 mM salt, respectively; "10" and "45" indicate temperatures of 10 and 45 C, respectively.)



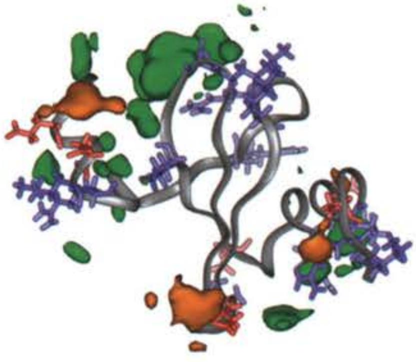
n10



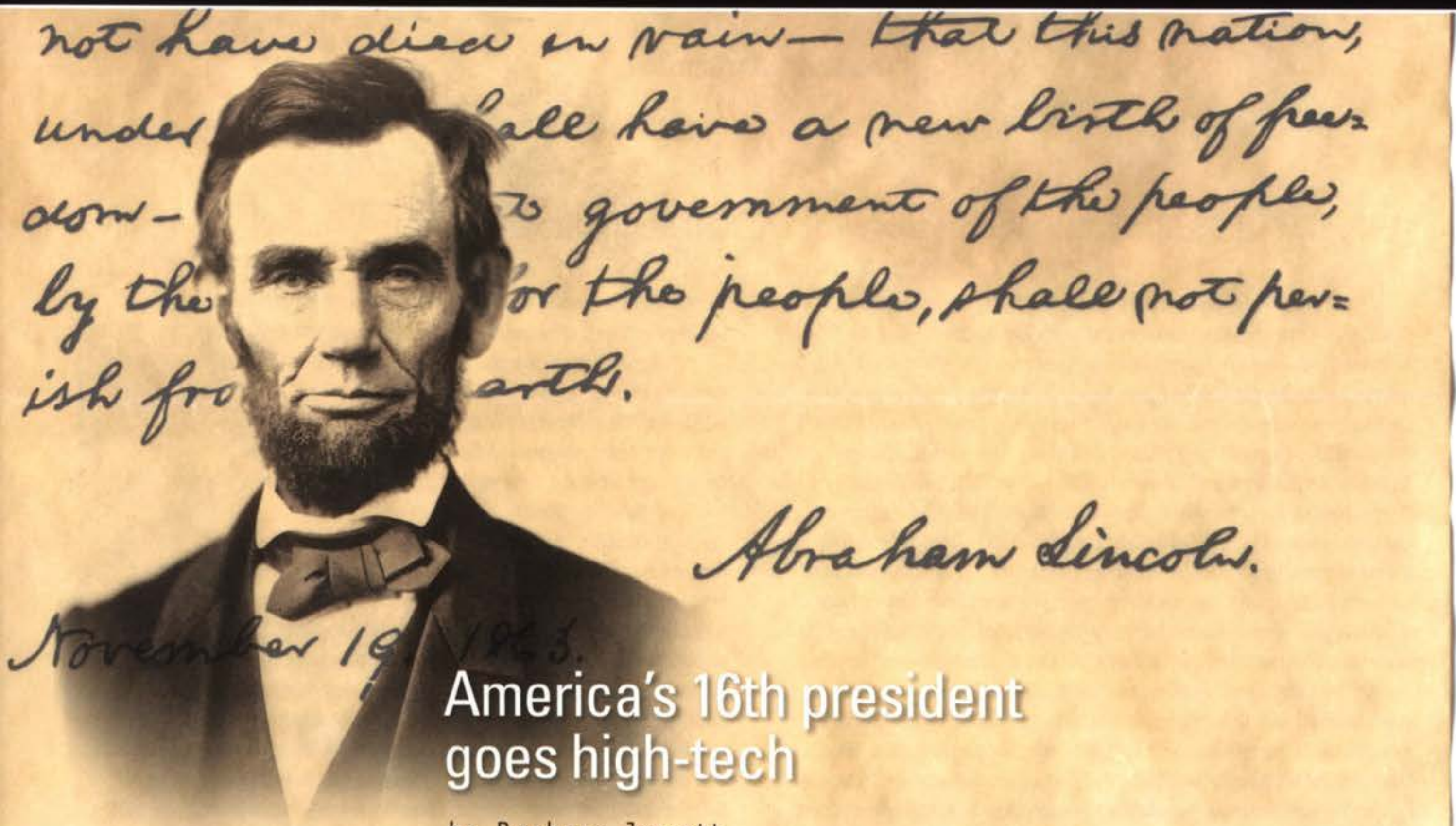
n45



s10



s45



America's 16th president goes high-tech

by Barbara Jewett

Using pen and ink to scribe thoughts on paper by the light of a candle, Abraham Lincoln surely never envisioned a world where news traveled around the globe via CNN's satellites or instant messages over the World Wide Web. But today Lincoln scholars worldwide have access to a life's worth of writings by America's 16th president, thanks to the state of Illinois and NCSA.

"The Papers of Abraham Lincoln is the most significant research initiative of the Abraham Lincoln Presidential Library and Museum," says Illinois State Historian Thomas F. Schwartz, interim executive director of the museum. "The project is now identifying all incoming and outgoing correspondence of Lincoln and capturing color images of these documents. The ultimate goal is to provide an online research and reference work that will provide the color images, transcriptions, and editorial matter for these historical documents, including Lincoln's extensive legal practice."

A project of the Illinois Historic Preservation Agency and the Abraham Lincoln Presidential Library and Museum, and co-sponsored by the University of Illinois at Springfield, the undertaking consists of three series: Lincoln's legal papers, his personal and political correspondence before he became president, and his presidential papers. Project staffers are traveling the country visiting archives searching out anything written by, or to, Lincoln. Items of questionable authenticity are verified by handwriting experts.

Once each series is complete and available to the public, users can view the images of the original documents or search the text of the transcriptions and annotations through a sophisticated search engine. "The first series, that of Lincoln's legal papers, will soon be online," says Daniel Stowell, project director and editor, offering access to more than 96,000 documents from Lincoln's legal career through the project's website.

"Our own internal storage capabilities were quickly reaching capacity with the bulk of the project still to be realized," says Schwartz. "Fortunately, Kirk Hard, NCSA's governmental relations officer, had followed our project and he served as the liaison that led to the wonderful collaboration we now enjoy with NCSA."

According to Michelle Butler, technical program manager for the storage enabling technology group, "NCSA is helping free up storage on servers in Springfield by providing a permanent storage archive. The archive server is providing two copies of the data in two different locations for safekeeping. Our group also has created tools to enable the storage process to be easier and require less management by adapting UberFTP to meet their needs." UberFTP, developed at NCSA, is the first interactive, GridFTP-enabled ftp software.

Once most of the document gathering is complete, says Butler, project staffers in Springfield will catalog the items, creating an online catalog with links to the items, the majority of which will be links to NCSA archives. "A person will visit the Papers of Abraham Lincoln website, search the catalog, click on the link to the item, and on their screen be viewing an item that is coming to them from NCSA servers." She notes that the most-requested items will probably remain on the project's servers in Springfield but be backed up at NCSA.

"NCSA's offer to host the archival images that we have collected from more than 100 repositories thus far is a vital part of our overall effort," notes Stowell.

For further information: http://www.papersofabrahamlincoln.org

IBM, NCSA collaborate on Cell applications

Last year NCSA researchers met with IBM's chief scientist/chief architect to hear an overview of Cell Broadband Engine™ architecture—design objectives, motivation, decisions in the design, and the computing paradigm shift from single thread performance to application-specific system customization. Since then, NCSA's Innovative Systems Lab (ISL) has begun working to port computing applications to Cell using IBM's Software Development Tool kit. These applications include MILC (quantum chromodynamics), NAMD (molecular dynamics), DEM (discrete element modeling of large numbers of particles), and PyQuante (a suite of programs for developing quantum chemistry methods).

Overall, results demonstrate a tremendous potential for scientific computations with Cell's multi-core implementation in an HPC environment. Porting of the MILC application is nearly complete. NCSA's Guochun Shi is working on a way to effectively exploit Cell to achieve 50 to 100 times faster performance than traditional HPC clusters. "We have high hopes for this architecture, but recognize that we may need to make significant modifications to our code to make the best use of this new hardware," says Steven Gottlieb, a physics professor at Indiana University.

"We need to either find an effective algorithm with existing memory layout in MILC, or change the entire MILC memory allocation to make it suited for SIMD in Cell," says Shi. Currently, there are eight simple SIMD (single instruction, multiple data) cores, containing a synergistic processing unit, local memory, and memory flow controller.

Exploring engineering and science applications with the first-generation Cell chip is a challenge, but NCSA and IBM researchers expect to find a path to supercomputing speeds. Cell has outstanding performance attributes, especially in multimedia applications (Cell's design is based on the Sony, Toshiba, and IBM partnership for the gaming industry), and the chip is applicable to a wide range of platforms. To further advance Cell, the Parallel Programming Lab in the University of Illinois at Urbana-Champaign's Computer Science Department is focusing on parallel programming models, compilers, and libraries.

NCSA and UIUC team on grain-tracking patent


Hornbaker

Two NCSA employees and a University of Illinois faculty member received a patent for a process to track grain from the point of harvest until the last moment before use. Robert Hornbaker, a professor of agricultural and consumer economics, and NCSA's Volodymyr Kindratenko and David Pointer developed a way to use RFID (radio-frequency identification) tags to provide grain elevator operators, food manufacturers and consumers with a wealth of information about the grain that enters our food supply, improving food processing efficiency and safety. Their collaboration was fostered by the NCSA/UIUC Faculty Fellowship Program; Hornbaker was a fellow in 2002–2003.

In the patented system, small RFID tags are distributed among the grain in the combine hopper at harvest; almost immediately the RFID tags can record the age, point of origin, and other information about the grain. Because the tags are approximately the same size and weight as individual pieces of grain, one or two tags simply float amid every 100 to 200 bushels of grain as it travels. Then, near the end of the process, the RFID tags can be removed from amid the grain, and the information stored on each tag can be read.

One of the keys to the patent is the way in which time is recorded: through an absolute atomic clock time/date stamp. The tag also stores the exact location of the event, using the worldwide latitude and longitude provided via a Global Positioning System inside.

Spend your summer at NCSA

Huerta

A notice on a listserve caught Timothy Huerta's attention and prompted him to apply for one of the first NCSA Summer Faculty Fellowships. A professor in the health organizational management program at Texas Tech University in Lubbock, Huerta spent this past summer working at NCSA developing a network-centric approach for tobacco control research. The goal of comprehensive tobacco control programs is to reduce disease, disability, and death related to smoking by preventing young people from starting, encouraging young people and adults who smoke to quit, eliminating nonsmokers' exposure to secondhand smoke, and identifying and eliminating the disparities related to tobacco use and its effects among different population groups.

"The summer fellowship provided me the opportunity to explore tools that enable researchers to work together, even though geographically separated. I looked at things from the non-technical perspective—what do people working in similar research areas need in a virtual community in order to interact—and NCSA provided the technical expertise to make it happen," says Huerta of the web portal that's being created. "People need a very simple way of managing their complex research relationships across university firewalls."

His work is an exemplar for public health, he says, and should be adaptable to other health concerns such as diabetes or heart attacks, allowing researchers and public health officials to look at problems holistically.

Huerta says he was already collaborating with Nosh Contractor, director of NCSA's Science of Networks in Communities research group, when he saw the summer fellowships announcement. Being a Summer Faculty Fellow, he says, allowed him to engage in a research project with national, possibly even global, implications that otherwise was not possible.

Building upon a successful program launch this year, funding will again be available for six summer fellowships in 2007 for faculty members who are not at the University of Illinois at Urbana-Champaign but interested in contributing to the creation of a national cyberinfrastructure for science and engineering. Fellowships include 10 weeks' salary, local housing expenses, and travel to and from Urbana-Champaign. Complete application details will be online, http://fellowships.ncsa.uiuc.edu, in early 2007.

NCSA a model for South African center

Calling the two months he spent as an International Fellow at NCSA "my best two months in academia," Peter McMahon becomes very animated as he talks about what he learned while in Urbana. During his time at the center he was able to study reconfigurable computing, both learning about it as a broad area and investigating the performance of algorithms that are of interest to scientific initiatives at the University of Cape Town, using the reconfigurable computers in NCSA's Innovative Systems Lab (ISL).

When setting goals for the new Centre for High Performance Computing (CHPC) at the University of Cape Town in South Africa, experts there turned to NCSA for guidance with hopes of emulating its success, inviting Radha Nandkumar, director of NCSA's Campus Relations and International Affiliations, to serve on their advisory council. She invited Michael Ingg of the UCT Electrical and Computer Engineering Department and his team to NCSA, which led to McMahon, a senior in electrical and computer engineering at UCT, gathering knowledge and experience with the ISL. He plans to help with the set up of a reconfigurable computing lab for the CHPC while pursuing a master's degree. CHPC will be acquiring hardware at the end of this year and begin establishing their center.

McMahon explored what types of problems are best suited to reconfigurable computing and what benefits one can expect from them. One of the algorithms he implemented that seems to be reasonably suited to reconfigurable computers is a Monte Carlo simulation.

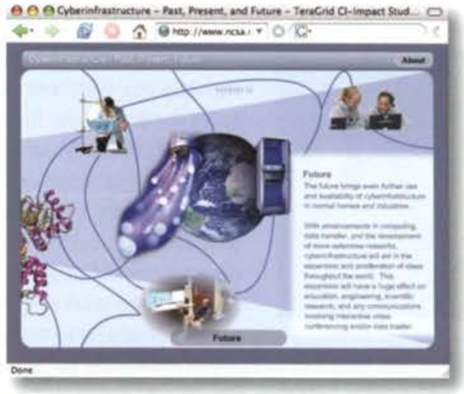
"Monte Carlo algorithms are a class of algorithms that are widely used in science to deal with problems that are difficult to solve deterministically when a large number of dimensions are involved," he explains. "One of the other algorithms I implemented—and had reasonable success with—was what is called a 'classical lattice gas,' which is a type of fluid dynamics simulation. That is, it simulates how a fluid (such as water or air) behaves, especially when there is an obstacle in its way, such as an aircraft wing."

Having access to the hardware and being able to use it was a very rewarding experience, he says. "But it was being able to work with ISL team members, including Craig Steffen, Volodymyr Kindratenko, and David Pointer, that made the visit a success. And the Reconfigurable Systems Summer Institute happening while I was here and being able to attend was a stroke of luck."

"The summer fellowship let me develop what health researchers need in a virtual community, and NCSA provided the technical expertise to make it happen."

Three reach finals of TeraGrid'06 student competition

Three students from NCSA were finalists in the student competitions at TeraGrid'06, held June 12-15 in Indianapolis. The contests showcased the talents and research capabilities of undergraduate and graduate students. The Cyberinfrastructure Impact contest asked students to create multimedia images and text designed to inspire others to learn about cyberinfrastructure, TeraGrid, and their impact in our daily lives while the Cyberinfrastructure Student Research contest asked students to present their own research results through posters.



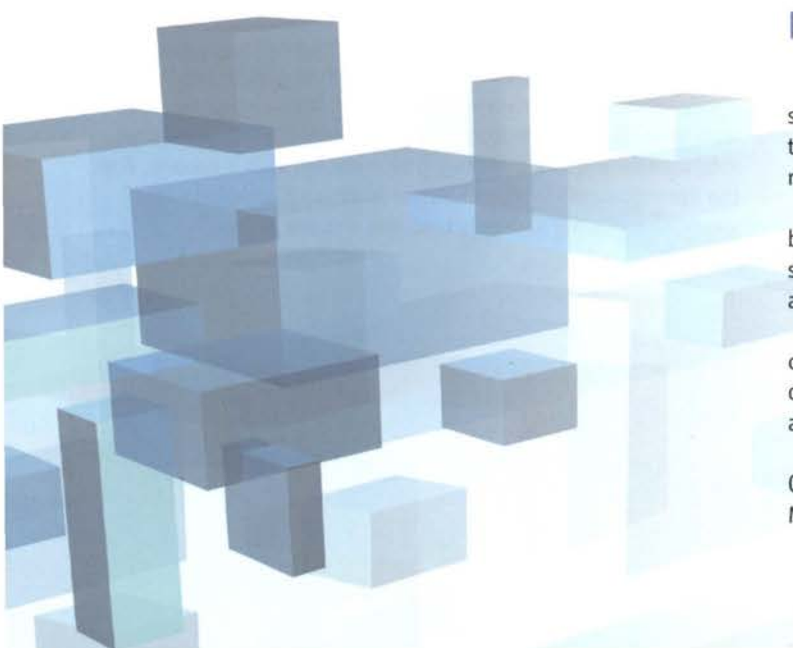
Cyberinfrastructure: Past, Present, and Future
August Knecht

Finalists from NCSA

August Knecht, a sophomore in computer engineering, serves as the webmaster for NCSA's Education Division, oversees the personal interface to the Grid (PIG) in the classroom program, supervises high school interns, and is developing a portal for the Engaging People in Cyberinfrastructure (EPIC) program. His entry described the past, present, and future of cyberinfrastructure; it can be seen online at http://www.ncsa.uiuc.edu/~augustk/teragrid.html.

Mike Mullan, a graduate student pursuing an MS in physics, is a former GK12 fellow and is working on a research project for the Computational Science Education Reference Desk to determine pathways to bring research to education communities. His submission to the student competition can be found at http://www.cornersuitestudios.com/mmullan/teragrid.

Undergraduate **Joel Poloney** was a student programmer in the Education Division; he assisted with the REVITALISE program, modifying Databridge with new features, including the ability to work with Excel documents, a new and improved interpolation algorithm, and a coordinate transformation algorithm. He also created extensive documentation on Databridge, including several tutorials and other guides on how to use the software. The latest release has been packaged with EasyViz 4.0 and has been named the NCSA Visualization Suite. His submission to the student competition focused on this work.



RSSI a success

Reconfigurable systems have the potential to dramatically speed computing and spur scientific discovery by enabling scientists to work with flexible "clay" that can be molded to fit their codes, rather than fixed "blocks" of computing power.

The Reconfigurable Systems Summer Institute (RSSI) hosted by NCSA July 10–14 provided a valuable opportunity for 80 domain scientists and technology developers from both industry and academia to come together to share information and ideas.

Presentations focused on the high-performance reconfigurable computing (HPRC) users' and programmers' points of view of development environments, compilers, libraries, and tools for HPRC, and included algorithms and applications.

RSSI was co-sponsored by NCSA, the Ohio Supercomputer Center, and the University of Manchester. Corporate sponsors were Mitrion, Nallatech, SGI, and SRC Computers.

Five in Focus

Petascale computing is now a realizable goal. But the path that will expose the full potential of petascale computers, allowing these machines to produce meaningful science and engineering results, is fraught with challenges. NCSA is working with partners around the country to overcome them. Here are five examples:

1 **System software.** Operating a computer of this scale requires compilers that translate friendly languages like Fortran into computer-ready assembly languages, checkpointing software that recovers a simulation run should the system crash during a calculation, data management software, and so forth.

2 **Applications.** These codes—which run simulations, analysis, and other important calculations—won't just automatically run on a system at this new level of scale. The codes will have to be re-imagined, and users will need intense support.

3 **Systems.** What hardware—and how much—is needed for a petaflops system is a crucial question. There's a need to understand what's imminent in the hardware area and how it will work.

4 **Storage.** A new scale of computing involves a new scale of archival storage.

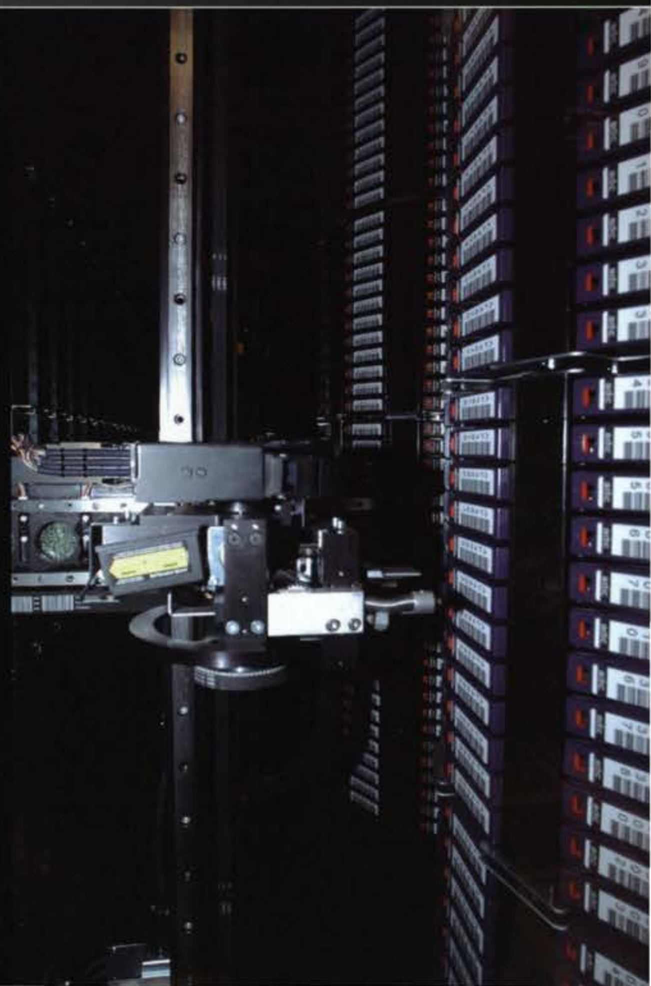
5 **Cyberenvironments.** NCSA's partnerships with research communities are the bridge that will bring together the new hardware, the new applications, and the researchers, allowing them to draw upon this extraordinary resource.

4 The right stuff

A petascale storage environment is full of interesting avenues for research, says Michelle Butler, technical program manager for NCSA's storage enabling technology group. For example, storage is the key to a balanced system so that the system can run without input/output (I/O) trouble and thus function efficiently during even the most taxing calculations. One area being explored is how much storage will be required for the number of petaflops so as to achieve balance. And the performance of the file system for the designated storage functions also needs to be examined, she notes.

Other areas that NCSA is exploring are how much storage will be required by applications running at petaflop speeds, and how to efficiently serve data from the archive environment. With a huge number of disk drives, she says, it is important to look at how the environment functions while keeping data integrity at the highest level; that is, not losing users' data owing to normal types of failures.

But all the research questions come down to this basic thought, says Butler: How do you keep a petascale machine's I/O wait at a minimal level? (I/O wait is when the CPU is waiting for some kind of storage system to reply about a request—read/write/fetch.) "The I/O wait needs to be kept at a minimum," she explains, "because a machine this powerful shouldn't be sitting idle waiting for data to show up." The bottom line: it is crucial to dedicate enough resources to facilitate the easy flow of information into and out of the machine. The trick is knowing how much is enough. But the combined efforts of NCSA and its partners will answer that question.



Galaxy formation

This image is a snapshot taken from a cosmology simulation by Brian O'Shea (Theoretical Astrophysics Group, Los Alamos National Laboratory) and Michael Norman (University of California, San Diego). The calculation follows the evolution of a region of the universe 250 million light years on a side, starting from only 16 million years after the Big Bang and ending at the present day, 13.7 billion years later. The purpose of this computation was to understand how a galaxy like our own forms and evolves over the lifetime of the universe. The calculation used a technique called "adaptive mesh refinement," which allows researchers to follow objects evolving on a large range of spatial and temporal scales while maximizing computational efficiency. Dark matter and the evolution of gas are followed separately, and a prescription for star formation and supernovae are also used.

This image shows two galaxies that are in the midst of a very complicated merger. Stars are shown as white streaks, and gas is shown in orange. When galaxies merge, they produce a halo of stars around them. Some will eventually settle back into the disk of the galaxy, and others will be jettisoned into intergalactic space. The "tidal tails" of gas are ejected from the galaxies during their closest point of approach and will eventually fall back onto the new galaxy.

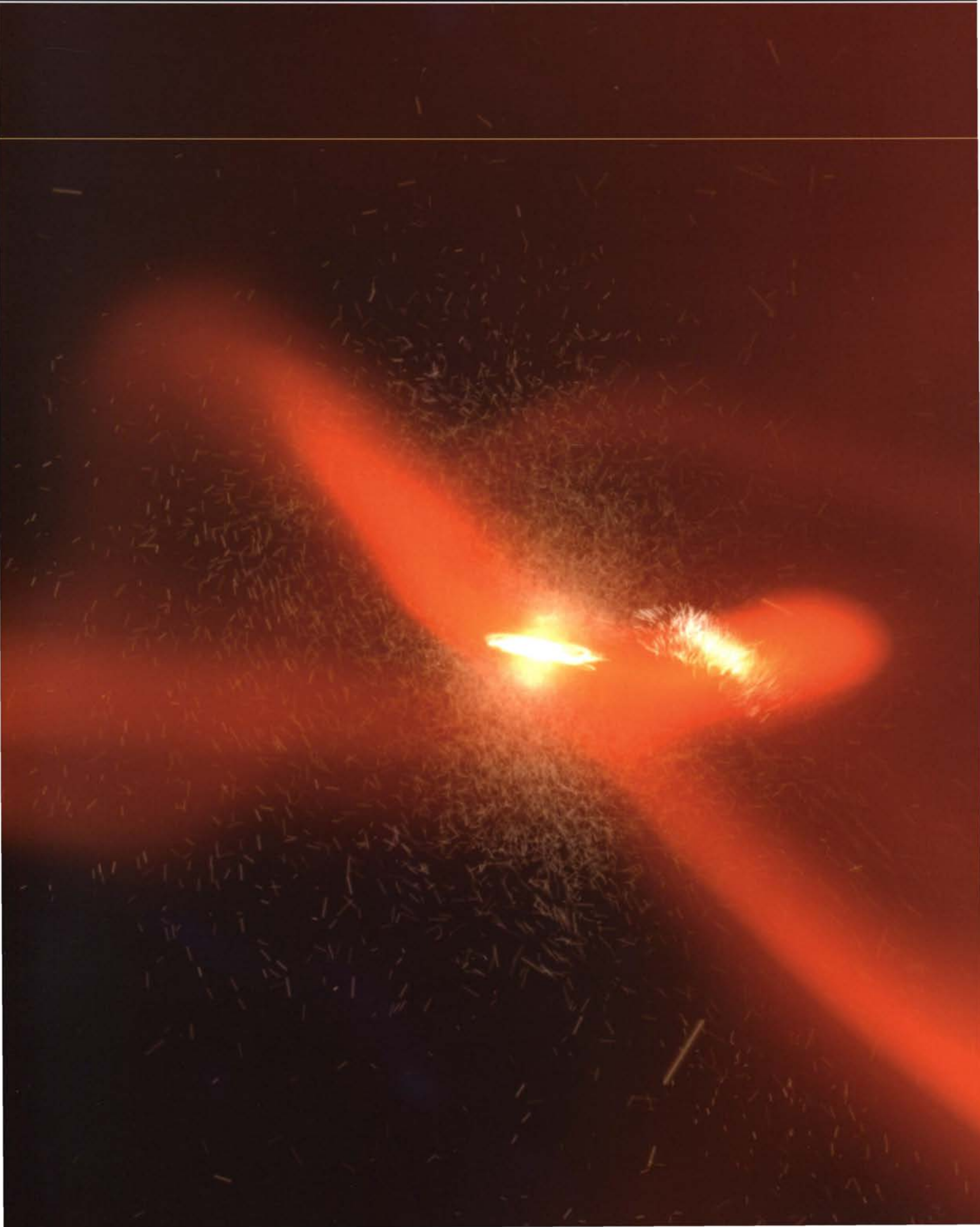
The image on the back cover shows the final stages of a merger of two galaxies into a larger one. A tidal tail of gas and stars (orange cloud and white streaks, respectively) is formed during the merger. This tail is similar to features seen by the Hubble Space Telescope in merging galaxies such as the Antennae Galaxies and will eventually fall back onto the galaxy.

This work was supported by NASA, the National Science Foundation, and the Department of Energy.

Team members

Scientific simulations
Brian O'Shea
Michael Norman

Visualization
Donna Cox
Lorne Leonard
Matt Hall
Stuart Levy
Robert Patterson



National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
1205 West Clark Street
Urbana, IL 61801

www.ncsa.uiuc.edu